OPTIMIZED CIRCUITRY FOR SENSOR INTERFACES: IN CMOS AND IN BRAINS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Ben Carl Johnson

January 2014

OPTIMIZED CIRCUITRY FOR SENSOR INTERFACES: IN CMOS AND IN BRAINS

Ben Carl Johnson, Ph.D.

Cornell University 2014

There is a need to optimize processing circuitry for sensor interfaces in both modern electronics and biology. Several analogs typically exist between sensing systems in electronics and biology: they have transducers to convert one type of energy or signal into one that can be processed by circuitry composed of transistors or neurons, they utilize an amplification stage to boost the sensed signal and suppress unwanted background signals, and then they have some means to store the signal or translate it into useful information. *Optimization* usually refers to power consumption, or, more specifically, optimizing the energy cost for processing a given amount of information. Circuit optimization requires selecting an appropriate architecture, bandwidth or speed, and output precision for a given task. In this dissertation I present several examples of circuitry in modern electronics and biology optimized for sensor interfaces.

Chapter 1 serves as an introduction to the dissertation. I discuss the similarities of electrical and biological sensor systems, how circuitry works to reduce unnecessary bandwidth and dynamic range of sensed signals for low power processing, and give an essential background to energy efficient CMOS amplifier design.

In Chapter 2 I present an orthogonal current reuse amplifier; a topology that circumvents the fundamental noise-power tradeoff in amplifiers by reusing bias current across independent amplifiers. This technique effectively increases am-

plifier g_m/I_D linearly with every additional amplifier at a small cost in headroom voltage.

Chapter 3 discusses the dynamics of gamma band oscillations recorded from olfactory bulb slices recorded with microelectrode arrays. Persistent gamma oscillations are induced in slice using methods previously reported for hippocampal slices and are shown to have multiple regions of coherent oscillatory activity across slice.

Chapter 4 presents microelectrode arrays developed in CMOS that scale to large electrode counts (+1000), have high spatiotemporal resolution (20kHz at 50μm pitch), and have integrated photosensors for correlating recorded electrical activity with optical stimuli.

In Chapter 5 I present a high-speed imager (> 1kfps) for calibrating MEMS intertial sensors in real-time. The imager utilizes polar symmetry to directly extract angular rotation information far more efficiently than standard, cartesian-based imagers.

BIOGRAPHICAL SKETCH

Ben Johnson earned his Bachelor of Science degree in Electrical Engineering from Oklahoma Christian University in 2007. He received his Masters of Engineering degree from Cornell University in 2008. In 2008 he joined the doctoral program in Electrical and Computer Engineering department at Cornell University.

Prior to graduate school, he worked at Martin Bionics (now Orthocare Innovations), a prosthetics research and development company, developing embedded control systems for prosthetic shoulder joints and ankles. While in graduate school, he interned with Analog Circuit Works creating inductively powered ICs for the Boston Retinal Implant Project, a collaboration of biologists, engineers, and physicians striving to restore vision. In the spring semester of 2013 he taught Digital VLSI Design at Cornell University, a senior-level course that teaches delay modeling, transistor level logic families, arithmetic circuits, and memories.

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor, Alyosha Molnar, who took a risk and gave me a chance when I needed something to do while my wife was attending vet school. He tolerated my sardonic and mediocre sense of humor and was committed to funding me while giving me the freedom to work on pretty much whatever I wanted. I've learned a lot from him: the nuanced workings of circuit design, effective teaching and management methods, and how to really be multidisciplinary. I also learned that there's a 2π everywhere except for switch-cap resistors, transmission lines really do reflect, and what I and Q actually mean. Most importantly he added useful words to my vocabulary such as oxygen thief, mental bandwidth, life tax, and PhD effect. I would also like to acknowledge my committee members, professors Thom Cleland and Amit Lal. Fortunately my graduate experience has been atypical in the sense that I've had several opportunities to interact and learn from my committee.

I have to acknowledge my fellow students from the Molnar Lab. Albert Wang for his help with ASP design in the first generation CMOS MEA and with whom I had several discussions about all sorts of crazy things (some ideas I naively tried to implement, but I ultimately decided not to document any of those "experiments" in this thesis), Caroline Andrews for including the stacked amp in her low-power radio thus making my research cooler, Changhyuk Lee for sticking with it and actually finishing the autorouter for the circular imager and testing the ladder rectifier (though I attribute my current coffee addiction to him – iced coffee is a gateway drug), and Sriram Sivaramakrishnan for being a catalyst in the circular imager testing process when I did not possess the required activation energy (we got the system working in less than a week, a new lab precedent). I also want to thank Shane Peace from the Computational

Physiology Lab with whom I've been able to do real neuroscience research. I think it's safe to say that by collaborating we've gotten further than either one of us would have alone.

Finally I would like to thank my wife, Katie, for her constant support and encouragement (e.g. baking cookies for my B Exam) throughout my graduate school career. I don't think there are many spouses who are understanding of late nights pushing rectangles around. I was never able to convince her to learn layout, though.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Signal Processing in CMOS and in the Brain

There is a need to optimize processing circuitry for sensor interfaces in both modern electronics and biology. Several analogs typically exist between sensing systems in electronics and biology: they have transducers to convert one type of energy or signal into one that can be processed by circuitry composed of transistors or neurons, they utilize an amplification stage to boost the sensed signal and suppress unwanted background signals, and then they have some means to store the signal or translate it into useful information. *Optimization* usually refers to power consumption, or, more specifically, optimizing the energy cost for processing a given amount of information. Circuit optimization requires selecting an appropriate architecture, bandwidth or speed, and output precision for a given task.

1.1.1 CMOS Sensors

A familiar example of optimizing a CMOS sensor interface is a radio receiver. The antenna converts electromagnetic radiation into electrical power. Received radio signals can have a dynamic range of about 96dB, which corresponds to a resolution of 16 bits if an ADC were connected directly to the antenna. Furthermore, the sampling rate of the ADC would need to exceed several GHz in order to do all the processing in the digital domain. An energy efficient 16bit ADC can achieve about 270fJ/conversion [11], but at a rate of 80MS/s. Assum-

ing this efficiency could be maintained at the 1 GS/s required for the receiver frontend, the ADC would consume 18W. For some perspective, the ADC itself would drain a cell phone battery in about half an hour (a typical cell phone power consumption is about 400mW, including everything: LCD, WiFi, etc.). Thus, a more appropriate architecture performs analog filtering up front, and down-converts the RF signal to baseband frequencies for more efficient digitization.

A similar argument could be made for extracellular neural recording instrumentation with high-impedance electrodes. While the signals of interest have very relaxed specifications relative to radio receivers (BW < 20kHz, SNR < 8 bits), building an area efficient SAR ADC to interface directly to the electrode would be difficult. For 8 bits over a 1mV scale, each quantization level would be $3.9\mu V$. To ensure that the system is not kT/C noise dominated, a total load capacitance of over 1nF would be required ($C = kT/(LSB/2)^2$). This corresponds to an area of about $1mm^2$ since modern CMOS processes typically have a capacitor density of $2fF/\mu m^2$.

Why is analog processing required up front for power efficiency? Digital design can be extremely low-power: if leakage is minimized, it consumes very little static power with large noise margins while analog circuitry requires a constant bias. This makes digital circuitry well-suited for storing information. However, the issue is converting from analog to digital with high-precision at high-speeds because of 2^N scaling factors, thus optimization requires converting at the lowest possible resolution and speed as possible. The function of analog pre-processing then is to reduce the bandwidth and dynamic range of signals. In the case of extracellular neural recording, low-pass filtering is required to

prevent aliasing thermal noise from the electrode and high-pass filtering is used to block relatively large, low-frequency electrochemical offsets.

Power efficient analog filtering can utilize non-dissipative devices (inductors and capacitors; their voltage is ideally orthogonal to their current making the power consumption zero) to passively filter signals. However, non-idealities of these devices and large areas can make them impractical for many designs. Power efficient active circuitry relies upon subthreshold biasing of transistors. The advantage of the subthreshold regime is that the transconductance per current is maximized *and* the voltage headroom is minimized. The caveat of the subthreshold transistors is that their exponential behavior makes them sensitive to transistor mismatch, temperature variations, and power supply noise.

1.1.2 Biological Sensors

Interestingly, ion channels also operate as Boltzmann exponential devices like subthreshold transistors, a similarity first noted by Carver Mead [65]. Biological sensors rely upon extremely power and area-efficient architectures that operate in noisy and unreliable environments. As the transistor is the building block of CMOS circuits, the neuron is the building block of neural circuits. Transistors use either electrons or holes as their primary charge carriers whereas neurons utilize the transport of sodium, potassium, chloride, or calcium ions. Generally, in comparison to CMOS, neurons are extremely power-efficient. They consume an estimated $10^{-11}J$ of energy per action potential [71], or an average power of 0.66nW per neuron while the human brain operates at 14.6W [73]. Comparing the efficiency of neural circuits to CMOS circuits is difficult, mostly due to the

fact that the brain does not operate as a Von Neumann architecture. Furthermore, neurons implement operations far more complex than single transistors or logic gates. Neurons may have thousands of input signals each with independent, graded levels of excitation and inhibition, which can then be encoded via different types of neurotransmitters. In fact, inhibitory and excitatory synapses may serve as better functional analogies to NFETs and PFETs. Within the neuron itself, discrete electrical and chemical compartments may carry out further graded, non-linear computations [68]. Tens of thousands of these compartments then interact in complicated ways to drive the "output" of a cell. Consider also that the synaptic connectivity can be bidirectional with built-in feedback and temporal plasticity.

Analogies between the architecture of biological and electrical sensory systems can be made and still provide valuable insight. In general, signals are transduced, amplified, high-pass filtered to decrease the dynamic range, low-pass filtered to eliminate unnecessary bandwidth, and discretized for enhanced noise margin or storage. For example, rods and cones of the retina absorb photons via a light-sensitive receptor protein (*transduction*), triggering a biochemical cascade (*amplification*). Furthermore, rods and cones have inherent shut-off mechanisms which act as feedback to return the cells to their resting state and adapt to continuous light levels (DC-rejection, *high-pass filtering*).

The olfactory system, which is revisited later in this thesis, has a different computational challenge compared to the auditory or visual sensory systems. The auditory system consists of a 1-D tonotopic map of frequencies. Visual stimuli are slightly more complex in that they must map two-dimensional space but the dimensionality needs are met by the two-dimensional, retinotopic organi-

zation of the retina. The olfactory system, however, must accommodate a significantly higher order of dimensionality that is limited to the two-dimensional limitations of olfactory bulb (OB) surface circuitry [15].

Odor stimuli are *transduced* via olfactory receptor (OR) proteins expressed on the cilia of primary olfactory sensory neurons (OSNs). There are roughly 1000 different olfactory receptors in the mouse olfactory system, which are independently encoded by a specific gene, that can respond to a myriad of odors [12]. All of the receptors belong to the mammalian G-protein-linked receptor superfamily of over 1000 genes [2]. In mice, about 20% of the OR genes are pseudogenes compared to about 60% in humans [32]. Even so, humans can distinguish over 10,000 distinct odors with a relatively reduced olfactory system.

Odor investigation results in patterns of activation on the primary OSNs, the axons of which converge onto spatially discrete glomeruli within the OB. Each glomerulus receives axonal projections from a population of OSNs expressing the same OR protein. In mice, the roughly 1000 distinct glomeruli are low-noise chemosensory *amplifiers* with graded levels of activation, indicating there may be at least 1000 different odorant dimensions. Furthermore, most bulbar neurons will extend their dendrites into only a single glomerulus, creating a glomerular column similar to the gross organization of the cortex. Unlike the other sensory systems with lower-order dimensionality, signal enhancement cannot be completely mediated by lateral inhibition due to the lack of receptive field similarity of neighboring glomeruli. The OB transforms the pattern of activation through decorrelation across similarly activated glomeruli and normalizing input activation patterns to levels within the dynamic range of central neural circuitry (*high-pass filtering*) [15].

1.2 An Introduction to Low-Noise, Low-Power Amplifier Design

The purpose of this section is to provide a simple overview of noise and energy-efficient amplifier design, starting from basic noise sources in transistors and emphasizing the major design strategies. MOSFETs exhibit two main types of noise (unwanted output signal variance): thermal noise (known as Johnson noise to experimental physicists and Nyquist noise to theoreticians) and flicker noise ($1/f$ noise).

1.2.1 Thermal Noise

Thermal noise in circuits is the result of the random motion of thermally agitated electrons (hence *thermal* noise first characterized by J. B. Johnson [48]). Interestingly, thermal noise in a conductor is not directly affected by the current since electron thermal velocities are about 10^3 times higher than electron drift velocities. Another feature of thermal noise is that its spectrum is independent of frequency for less than 80THz [56]. The most significant source of thermal noise in MOSFETs is generated in the channel. The Norton equivalent can be modeled as a current source connected between the drain and source terminals with a power spectral density (PSD, in A^2/Hz) of

$$\overline{I_{o,th}^2} = \frac{4kT}{R_{CH}}\Delta f = 4kT\gamma g_m \Delta f, \tag{1.1}$$

where the overline indicates averaging and the effective channel resistance (R_{CH})

is $1/(\gamma g_m)$. The channel is not homogeneous as it conducts well on the source side and pinches off toward the drain. γ is typically modeled as $2/3$ for long channel devices. Velocity saturation increases γ for shorter channel devices and is roughly 1 for 180nm CMOS [37]. Also note that the ohmic contacts of MOSFETs (source/drain diffusions and the gate) also contribute noise, but can be reduced by good layout technique. Furthermore, unlike in high-speed RF circuits, the impedances in low-power analog design are often very high allowing designers to neglect thermal noise from the terminal resistances. The transistor r_o is actually noiseless because its a modeling parameter to account for finite dI_D/dV_{DS}.

1.2.2 Flicker Noise

Flicker noise, however, cannot be neglected in low-power analog design since it is dominant at lower frequencies. Flicker noise is also called $1/f$ noise because its PSD exhibits a behavior inversely proportional to frequency. While flicker noise is ubiquitous in physical systems, it has no universally known mechanism. Flicker noise in MOSFETs primarily arises from the random capture and release of charge carriers by traps in the gate dielectric. Traps are caused by imperfections in the lattice structure, resulting in spatially localized levels in the energy band structure available for carriers to occupy. The characteristic $1/f$ spectrum is generated by the summation of many different trap spectra which individually exhibit a $1/f^2$ spectrum. MOSFET flicker noise is often modeled by a voltage source in series with the gate with a PSD (V^2/Hz) of

$$\overline{V_{in,f}^2} = \frac{K_f}{WLC_{OX}} \cdot \frac{\Delta f}{f}, \tag{1.2}$$

where K_f is a process-dependent parameter. For 180nm, $K_{f,NMOS} \approx 0.5 \cdot 10^{-25} V^2 F$ and $K_{f,NMOS} \approx 0.25 \cdot 10^{-25} V^2 F$. A plethora of expressions for flicker noise are in use and can quickly become quite complex and cumbersome. The critical points – aside from $1/f$ frequency dependence – is that flicker noise is a function of transistor area and not DC bias conditions, that PFETs *typically* have some advantage over NFETs (the channel in PFETs is "'buried"' further away from the oxide-silicon interface, but JFETs are best), and flicker noise improves with newer processes. Thus a larger area FET in a modern process little $1/f$ noise.

1.2.3 Common Source Noise

Consider the common source amplifier in Figure 1.1. The key observation is that when the noise is input-referred, the noise density is minimized when the input g_{M1} is maximized and the mirror g_{M2} is minimized for a given bias current, I_D. Note that the input-referred noise density is not dependent on output impedance (the total integrated noise is, however) and that the input device attenuates the noise contributions of the load device by a factor of $1/{g_{M1}}^2$. The EKV model (a region-continuous alternative to square-law, from **E**nz, **K**rummenacher, and **V**ittoz [22]) shows that g_M/I_D is maximized when devices are biased in subthreshold and minimized when devices are biased in deep saturation. Input devices should have a large W/L ratio to drive them into subthreshold and mirror devices should have a small W/L ratio. Increasing V_{OD} is another alternative to decreasing g_M, but be mindful that this increases the

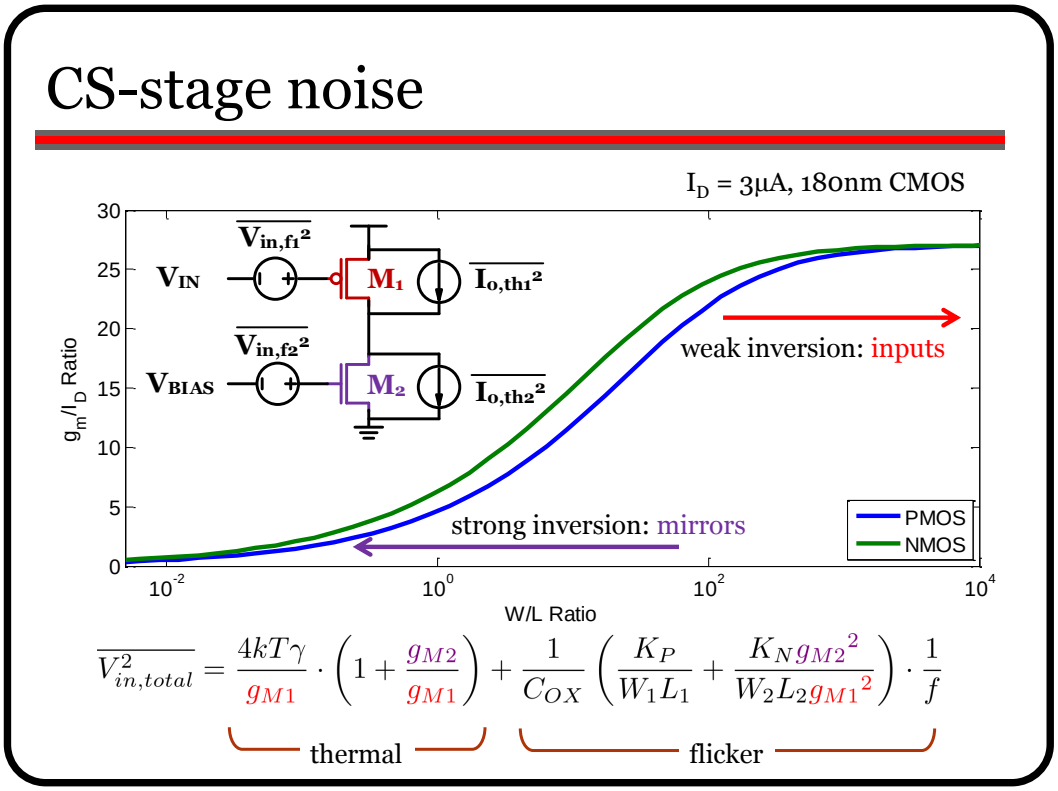

Figure 1.1: Noise minimization for a common source amplifier

overall power consumption (in saturation $g_M = 2I_D/V_{OD}$). In summary, after optimizing transistor biasing and ratios, noise density is decreased at the expense of either increasing transistor area or increasing the power consumption.

1.2.4 Source Degenerated Noise

Is source degenerating devices with resistors more power efficient than increasing V_{OD}? Firstly we examine source degenerating the input device with a resistor and compare its total thermal noise power to a simple common source amplifier assuming both circuits consume equivalent power and the active load has no noise contributions (Fig. 1.2). This case is quite simple: since subthreshold

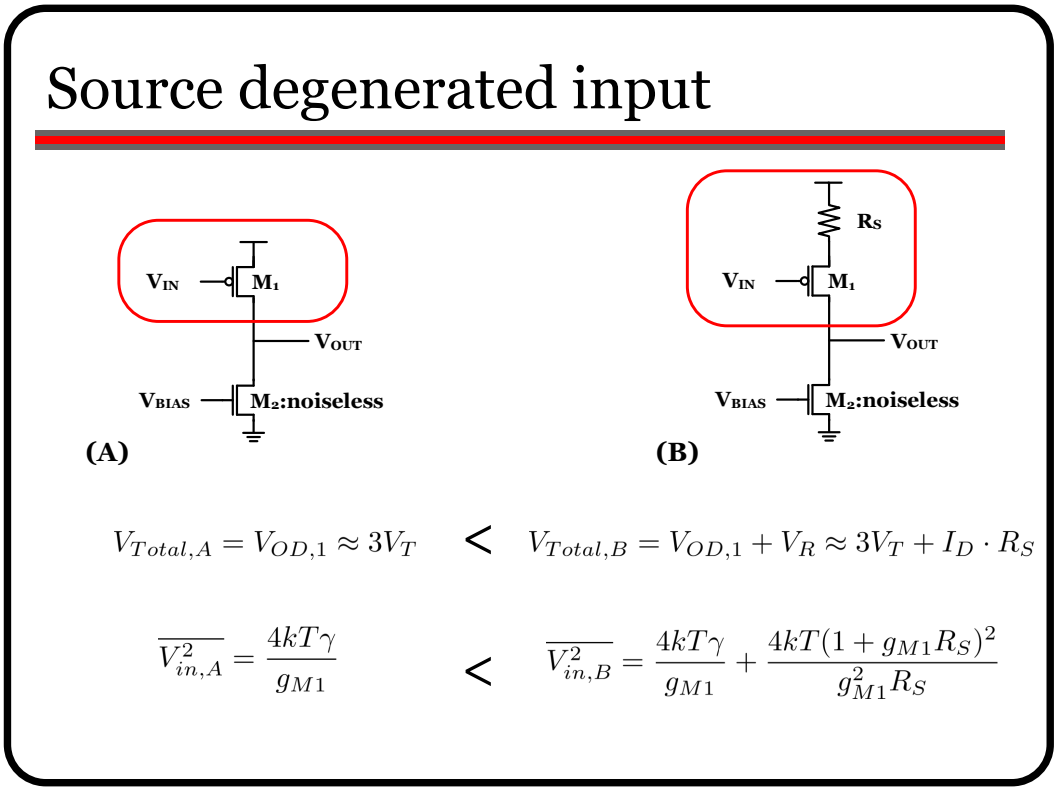

Figure 1.2: Noise of a source degenerated common source amplifier

devices require very little overdrive voltage that's independent of bias current (a few V_T for I_D to be V_{DS} independent), it's not power efficient to resistively degenerate. To add insult to injury: not only does source degenerating consume more power, it also has higher overall input-referred noise. Recall that power efficiency was achieved by maximizing the g_M/I_D ratio of the input device and source degenerating attenuates the transconductance. Note that the output noise may be smaller, but this is because the gain is decreased.

Secondly we examine the case of resistively source degenerating the active load (Fig. 1.3). Similar to before, we compare this structure to a standard common source amplifier and assume that both structures have the same power consumption. Intuitively this is a favorable option since decreasing the

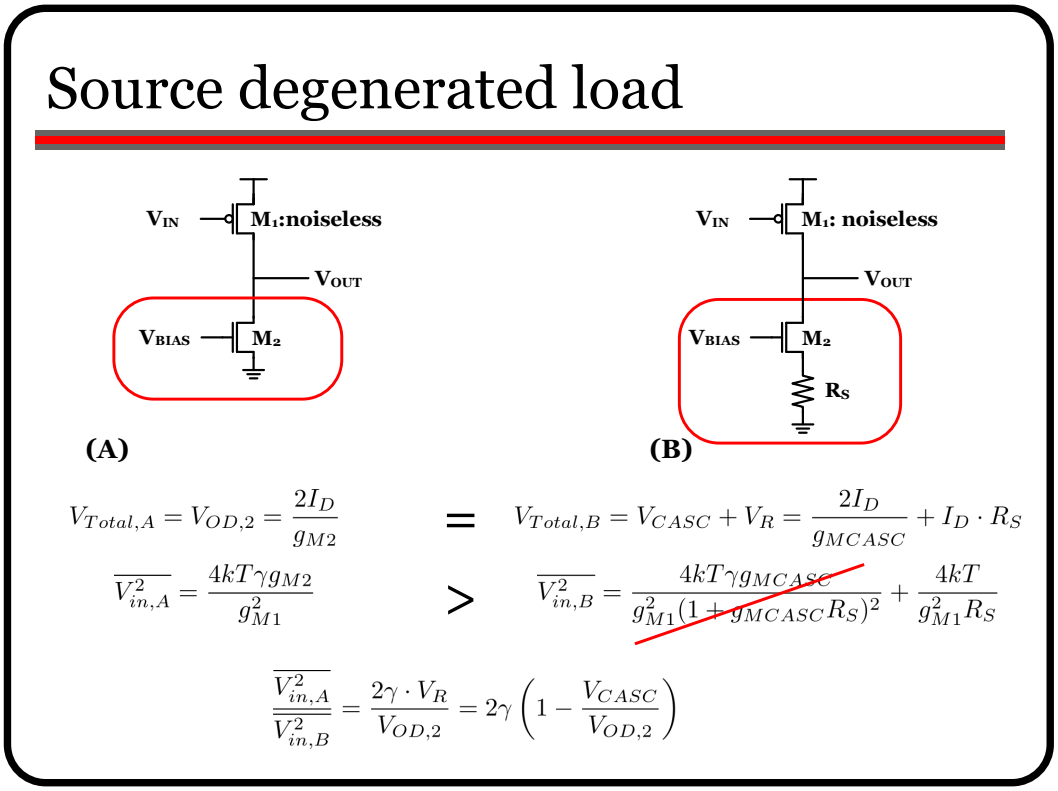

Figure 1.3: Noise of a source degenerated active load

transconductance and lowering the gain is desired for the mirror device. If we assume that the voltage across the resistor (V_R) is larger than the cascode device overdrive (V_{CASC}), then we can safely neglect the cascode's noise power contributions since they are attenuated by $1/(1 + 2V_R/V_{CASC})^2$, or more traditionally $1/(1 + g_{MCASC}R_S)^2$. Assuming that $V_{OD2} = V_R + V_{CASC}$, we find that the input-referred noise ratio between the two structures is

$$\frac{\overline{V_{in,A}^2}}{\overline{V_{in,B}^2}} = 2\gamma\left(1 - \frac{V_{CASC}}{V_{OD2}}\right), \tag{1.3}$$

implying that we see more of an improvement as we maximize V_R and minimize V_{CASC}. To put it another way, when we say that we want to minimize the

mirror's g_M for efficiency, we're essentially saying that we want to increase R_{CH}. What equation 1.3 effectively tells us is how much bigger R_S is than R_{CH} for the same headroom.

There are, however, a few caveats from this analysis. This analysis assumes that there is enough headroom to bias *both* the cascode and the resistor, which may be impractical for small supply voltages. It also assumes that we need the cascode device for biasing – the input device could be used as a subthreshold current mirror. Moreover, degeneration may see limited return on investment if the input device already dominates the noise performance. There are a few other useful observations, too. An advantage of resistive degeneration is that resistors typically exhibit less flicker noise than MOSFETs. Degenerated current mirrors are typically better matched (decreased g_M dependence). Furthermore, noise from the current mirror reference is attenuated. Firstly, the noise from the reference is degenerated. Secondly, noise generated by the reference is presented to the input of the amplifier bias, which is source degenerated also, decreasing its gain.

CHAPTER 2

OPTIMIZING NOISE AND POWER TRADEOFFS: ORTHOGONAL CURRENT-REUSE AMPLIFIER

2.1 Introduction to Power-Efficient Amplifiers

There is a fundamental tradeoff between power consumption and noise performance in amplifier design. Minimizing this tradeoff is critical for many applications, and has seen particular attention in the context of neural amplifier design. Significant effort has been made to optimize the power and noise tradeoff, for implantable devices [77], neural prosthetics [81], and microelectrode arrays [27], [47], [31]. A majority of neural applications require multiple recording channels to simultaneously record electric field activity across multiple sites in tissue. Increasing the number of recording channels allows for a greater number of single cell spike recordings and spatial analysis of local field potentials (LFP). However, extracellular instrumentation requires good noise performance to detect neural signals with microvolt amplitude. Additionally, the kinetics of neural tissue are sensitive to temperature changes, so heat dissipation (and so, power consumption) must be minimized. In reported recording systems, the frontend amplifiers dictate the overall noise and power performance of the system [39]. State-of-the-art energy efficient amplifiers attempt to minimize the power and thermal noise tradeoff by utilizing a combination of deep subthreshold design for maximum g_m/I ratio [38], folded cascodes with low current folds [87], reference sharing [64], or g_m-boosting via current-splitting [72] (see Fig. 2.1). Additional g_m-boosting techniques include complementary device inputs [41] or driving the back-gate [13]. Neural signals have spectral energy limited to rel-

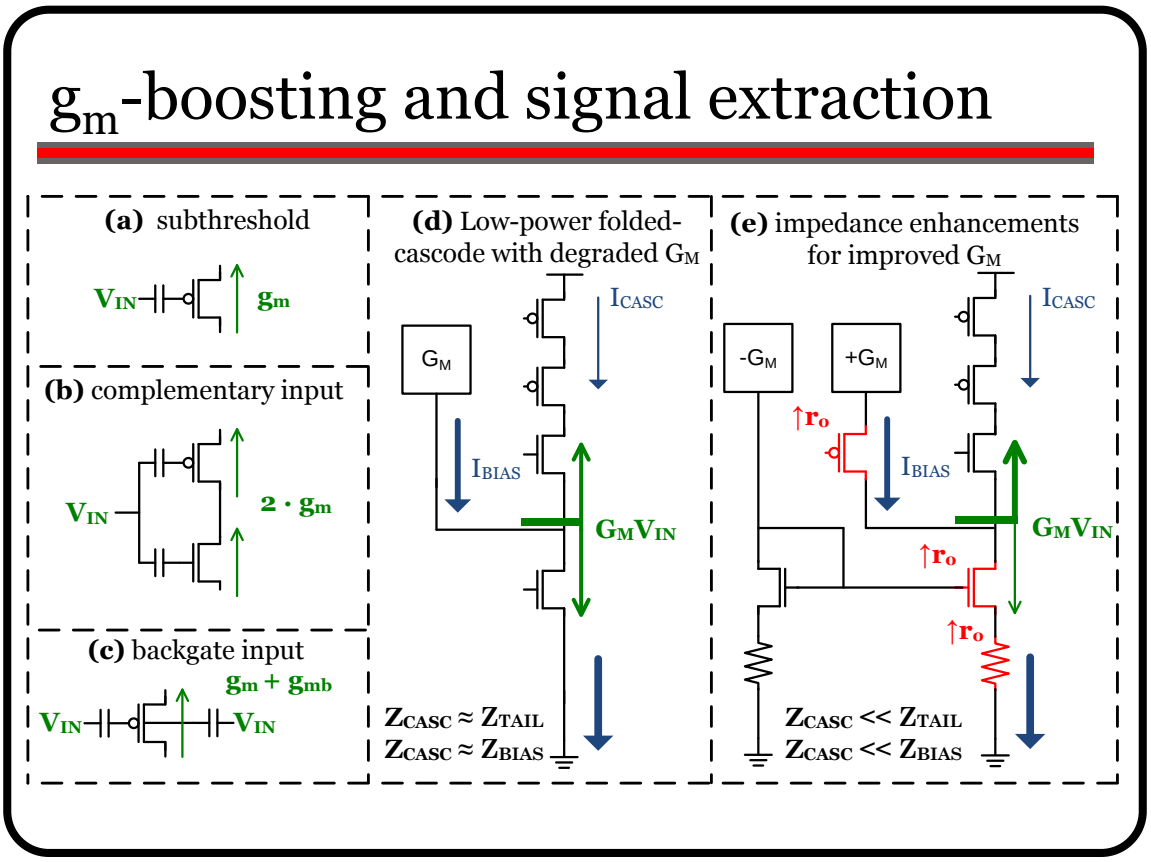

Figure 2.1: Examples of g_m/I-enhancement techniques

atively low frequencies, ranging from 300Hz to 6kHz for action potentials and less than 300Hz for LFP. Thus energy efficient amplifiers must also optimize for intrinsic $1/f$ noise by using large input devices [39] or chopping [19].

To date, the main goal of energy efficient design has been to maximize the amplifier's transconductance for a given bias current. A key point driving neural amplifier design is the observation that in a well designed amplifier, noise performance is dominated by the amplifier's input differential pair. Thus, the majority of bias current should be consumed by the input pair while subsequent circuitry can be biased at a much lower current [87]. An additional, important

point is that the noise of the amplifier is only weakly dependent on supply voltage. Since the amplitudes of action potentials and LFP typically range from 10μV to 1mV for extracellular recording, linearity requirements are relaxed relative to other applications, such as baseband amplifiers in high performance RF systems. Since the voltage headroom required by an input pair biased in subthreshold is minimal, amplifiers using stacked, self-biased subthreshold inputs can operate with low voltage supplies without compromising noise performance [41].

Stacked, current-reusing topologies have been used extensively in RF transceiver circuits to decrease power consumption, typically sharing current between the LNA and mixers [94], the oscillators and amplifiers [66], or the LNA, mixer and VCO [76]. These designs employ capacitors to AC-couple signals and decouple virtual grounds, allowing multiple RF or IF circuits to share the same DC voltage stack while reusing bias current. Additional voltage headroom is conserved by using inductive loads rather than resistive loads. However, these techniques translate poorly to the low frequency regime, where decoupling of AC grounds is much harder and inductors are impractical. In this paper, we introduce a low-frequency-compatible technique that avoids these problems by stacking differential input pairs to reuse bias current between channels while preserving linear independence between signals. Additionally, orthogonal current-reuse is compatible with several other energy efficient techniques, such as deep subthreshold design [38] and low-power folded cascodes [87].

This chapter is organized as follows. Section 2.2 discusses metrics used to evaluate the noise and power tradeoff in amplifiers and introduces orthogonal

current-reuse as a means to minimize the noise and power tradeoff in amplifiers. Section 2.3 discusses the design of the implemented four-channel amplifier. Section 2.4 presents test results of the fabricated chips and neural data acquired from a mouse olfactory bulb slice. Section 2.5 concludes the chapter with the introduction of a new amplifier metric, further noise improvements using chopping modulation, and two implementations of the orthogonal current-reuse amplifier with a wireless rectifier and low-power receiver.

2.2 Orthogonal Current-Reuse

2.2.1 Noise Efficiency Factor

A common figure of merit for noise and power constrained amplifier performance is the noise efficiency factor (NEF), introduced in [78]. NEF quantifies the tradeoff between noise, bias current, and bandwidth in amplifiers with a first-order roll-off and is given by:

$$NEF = V_{rms,in}\sqrt{\frac{2 \cdot I_{total}}{\pi \cdot V_T \cdot 4kT \cdot BW}} \tag{2.1}$$

where I_{total} is the current consumed by the amplifier, BW is the 3-dB bandwidth in hertz, and $V_{rms,in}$ is the input-referred noise voltage. NEF is the ratio of the input-referred amplifier noise to that of an ideal BJT common emitter amplifier biased with the same supply current. Practical amplifiers always have NEF > 1 since they consist of more than one noise source. Most practical amplifiers have a differential topology, implying a NEF of at least 1.4. Additionally, the perfor-

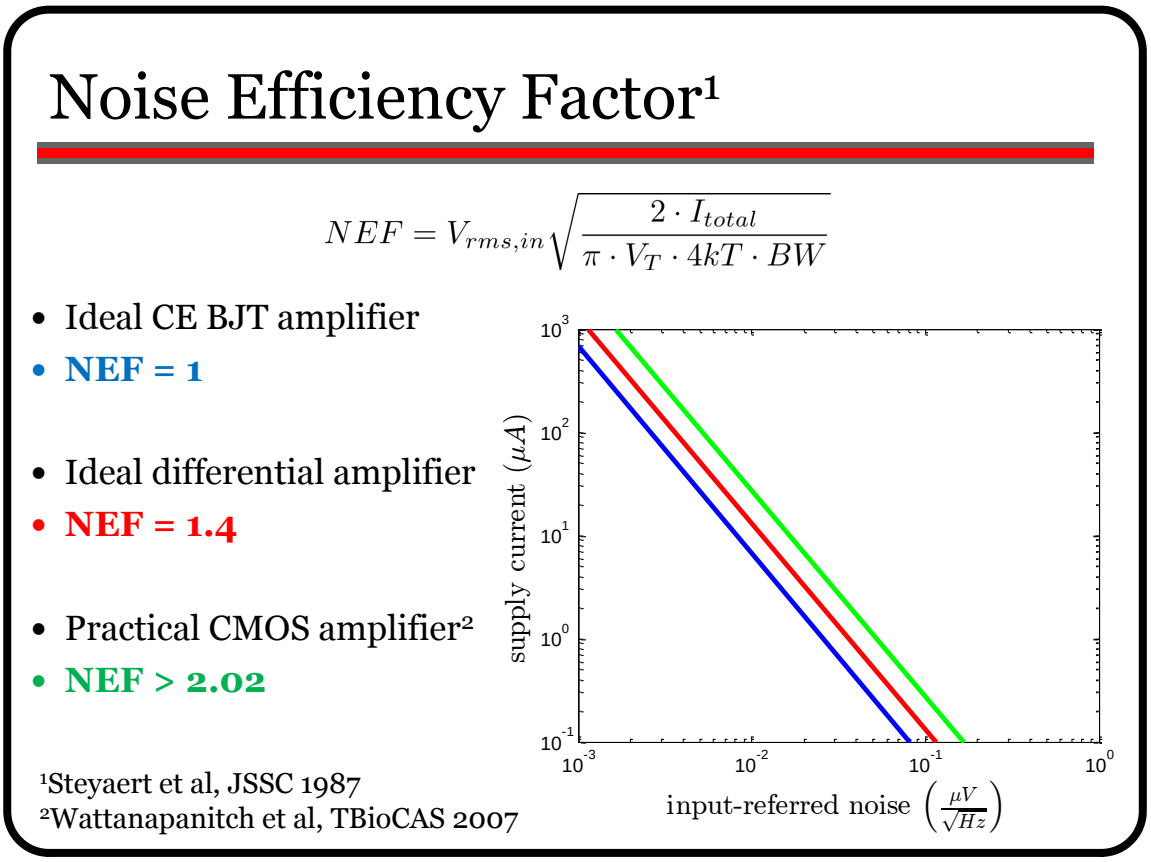

Figure 2.2: Power vs. Noise: Noise Efficiency Factor (NEF)

mance of CMOS devices is further degraded by a reduced subthreshold slope relative to a BJT (Fig. 2.2). While they contribute little to overall noise, active loads and folded cascodes consume current, further increasing NEF. Nonetheless, various techniques have allowed state-of-the-art CMOS differential amplifiers to achieve a NEF under 3.0 [87], [93].

While NEF is a valuable figure of merit for noise and power constrained amplifier design, it more accurately reflects current efficiency rather than power efficiency. A low NEF does not, therefore, necessarily indicate low power consumption since it does not account for supply voltage. NEF is minimized when

the input devices of an amplifier are biased in subthreshold to maximize their g_m/I_D, while the active-load bias transistors are biased in extreme saturation to minimize their g_m/I_D. Thus, an active-load device should be biased with a large overdrive voltage to provide better matching and lower noise. Therefore, when minimizing noise, input devices define the required bias current, but active-load transistors primarily define the required supply voltage and generally consume more power than the input devices.

Thus, design of a power optimized, high-gain, low-noise amplifier requires maximizing the proportion of current through the input differential pair while mitigating the overhead power associated with biasing the input and maintaining sufficient headroom for the output to swing a significant voltage. We introduce orthogonal current-reuse as a technique that simultaneously meets these requirements [42], [44].

2.2.2 Orthogonal Current-Reuse

Orthogonal current-reuse employs input stacking to increase the overall g_m of a differential amplifier proportional to the number of layers of stacked differential input pairs. Since each differential pair only requires one V_{DS} of voltage headroom, the increase in g_m comes at a minimal cost in voltage. Orthogonal current-reuse in a two-channel design requires that the second channel's input be split into two identical, half-sized differential pairs that use the drain currents from the first channel as tail currents, as shown in Figure 2.3. The differential pairs of channel two are independent yet effectively in parallel since they have common inputs. Additionally, the DC bias current is split evenly between them

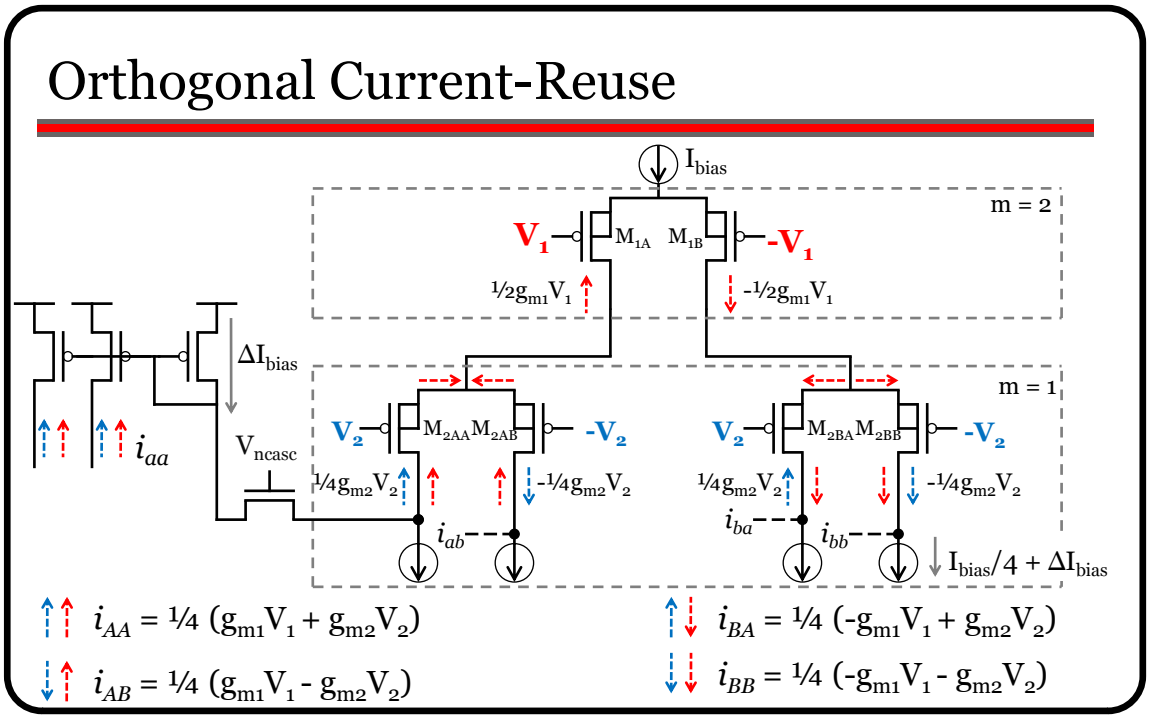

Figure 2.3: Two-channel example of orthogonal current-reuse

($I_{bias}/2$), meaning the g_m's of both pairs are equivalent. Since the total bias current through channel one is equal to the sum of the current through both pairs of channel two, the overall g_m's of both channels are equivalent. The effective widths of both channels are made identical to ensure similar operating points and noise performance. The active loads at the bottom of the stack are designed to sink one-quarter of the total bias current, amortizing the power consumption of the mirror devices across multiple channels without compromising noise performance. Orthogonal current-reuse can accommodate multiple signal channels each with the same amplifier g_m by stacking more differential pairs, further increasing the amplifiers' efficiency.

The small signal output currents from channel one are present on the outputs

of channel two as shown in Figure 2.3. For example, like the DC bias current, the small signal drain current from transistor M_{1A} will be split evenly between transistors M_{2AA} and M_{2AB} of channel two. Thus, the small signal output at each leg of the stack encodes a unique combination of a V_1 and V_2 component. This can be expressed in matrix form:

$$\vec{\boldsymbol{i}} = \begin{bmatrix} i_{aa} \\ i_{ab} \\ i_{ba} \\ i_{bb} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \cdot \frac{g_{m1}V_1}{4} + \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \cdot \frac{g_{m2}V_2}{4} = \frac{G_m}{4} \cdot \boldsymbol{A} \cdot \vec{\boldsymbol{V}}_{in} \tag{2.2}$$

where the general case is $g_{m1} = g_{m2} = G_m$ and, for example, i_{ab} is the small signal output current of transistor M_{2AB}. Since the dot product of the two column vectors of $\boldsymbol{A}$ is zero (they are orthogonal), the original voltage inputs can be independently reconstructed from these currents. Adding currents i_{aa} ($i_{aa} = 0.25g_{m1}V_1 + 0.25g_{m2}V_2$) and i_{ab} ($i_{ab} = 0.25g_{m1}V_1 - 0.25g_{m2}V_2$), for instance, will sum the V_1 components and cancel the V_2 components, including noise from M_2, resulting in a small signal current equal to $0.5g_{m1}V_1$.

In order to perform the required recombination to recover V_1 and V_2, a folded cascode is used to fold the output currents for recombination (Figure 2.3). The cascode is biased at a much smaller current than the input stack to conserve power ($\Delta I_{bias} << I_{bias}$). Each output current is mirrored twice and summed with the appropriate polarity across a load to construct output voltages. This operation can also be expressed as a matrix:

$$\vec{\boldsymbol{V}}_{out} = \begin{bmatrix} V_{1out} \\ V_{2out} \end{bmatrix} \approx R_o \cdot \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} i_{aa} \\ i_{ab} \\ i_{ba} \\ i_{bb} \end{bmatrix} = R_o \cdot \boldsymbol{A}^{\boldsymbol{T}} \cdot \vec{\boldsymbol{i}} \qquad (2.3)$$

where R_o is the output load and V_{1out} and V_{2out} represent the differential amplified voltages of inputs V_1 and V_2. Note that combining the matrices from equations (2.2) and (2.3), the matrix product $\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A}$, yields 4 times the identity matrix (because the columns of A are orthogonal), effectively gaining up each channel independently:

$$\vec{\boldsymbol{V}}_{out} = \frac{G_m R_o}{4} \cdot \boldsymbol{A}^{\boldsymbol{T}} \cdot \boldsymbol{A} \cdot \vec{\boldsymbol{V}}_{in} = G_m R_o \cdot \vec{\boldsymbol{V}}_{in}. \qquad (2.4)$$

This technique is extendable to more than two channels. An *n*-channel orthogonal current-reuse amplifier with *n* layers generates 2^n output currents, each of which is mirrored *n* times in order to construct output voltages. Figure 2.4 shows the general form of an amplifier with *n* layers with a total current consumption of $I_{bias} + (n+1)2^n \Delta I_{bias}$. Assuming sufficient headroom, the per-channel NEF of a *n*-stack orthogonal current-reuse amplifier will roughly decrease as $\sqrt{n}$ as long as $I_{bias} >> (n+1)2^n \Delta I_{bias}$. Orthogonality of output currents is necessary in order to benefit from averaging the output currents, otherwise the signal and noise from each input differential pair will degrade multiple channels.

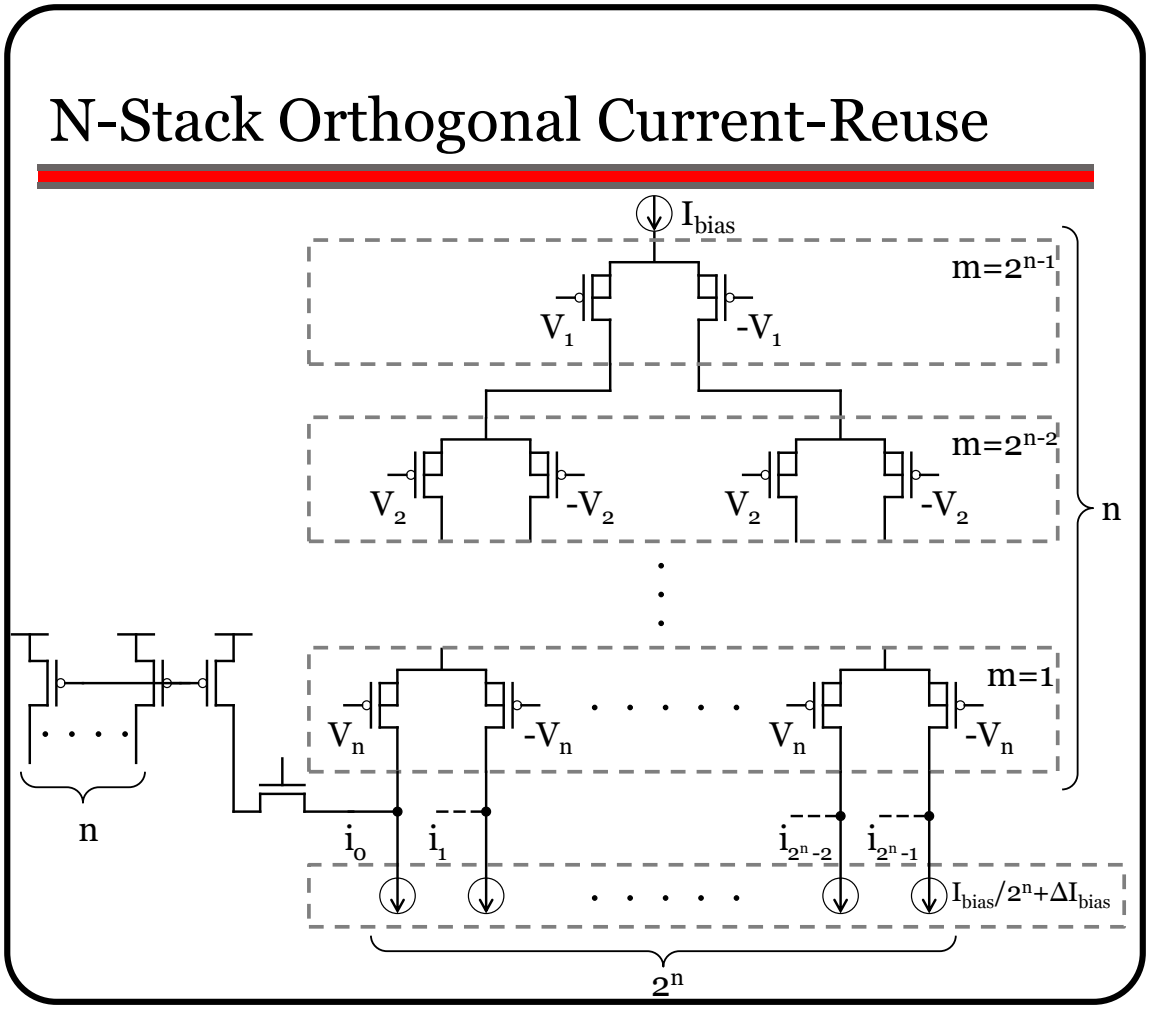

Figure 2.4: *N*-channel example of orthogonal current-reuse

2.2.3 Nonideal Orthogonality

One reasonable concern with orthogonal stacking is that it bears a topological similarity to a Gilbert mixer, and may be expected to generate significant cross-channel nonlinearity. In reality, the signal from the first channel modulates the g_m of the second channel. However, the configuration of the orthogonal stack and associated recombination circuitry acts to suppress cross multi-

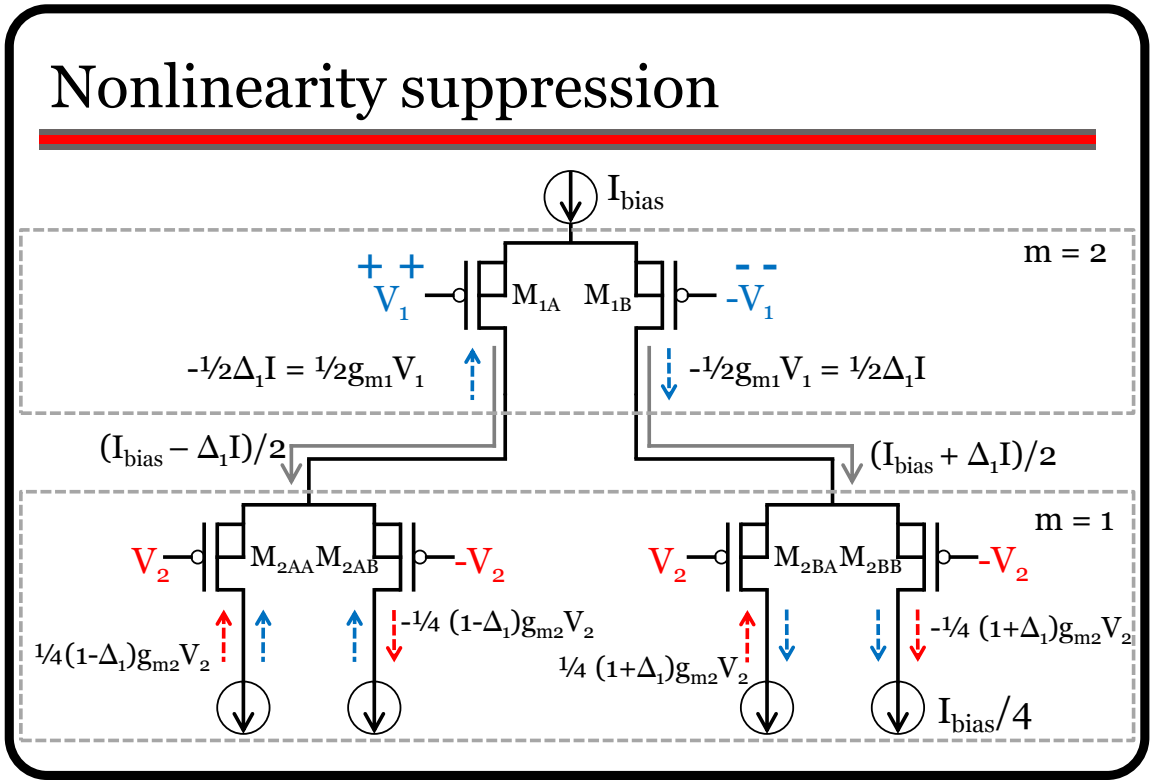

Figure 2.5: Suppression of nonlinearities with orthogonal current-reuse

plication terms. Equation (2.5) is the modified gain matrix including the cross multiplication terms for a 2-channel stack:

$$\begin{bmatrix} i_{aa} \\ i_{ab} \\ i_{ba} \\ i_{bb} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \cdot \frac{g_{m1} V_1}{4} + \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \cdot \frac{g_{m2} V_2}{4} + \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \cdot \frac{g_{m1} V_1 V_2}{4 \eta V_T}. \tag{2.5}$$

When the input devices are perfectly matched, the crosstalk component (the third vector) is orthogonal to the columns of $\boldsymbol{A}$ and is suppressed during recombination. To see this, suppose a large input signal is applied differentially to

input V_1 such that the signal current from channel one is large enough to affect the bias current of channel two as shown in Figure 2.5. The signal current increases the g_m of subthreshold transistors M_{2BB} and M_{2BA} by $g_{m1}V_1/4\eta V_T$ (where η is the subthreshold ideality factor), while decreasing the g_m of the other differential pair by the same amount. The differential topology ensures the overall g_m of channel two remains constant, compensating for signal, noise and offsets from the first channel. Crosstalk can occur, however, when there is mismatch among the input devices, specifically between the two differential pairs of the second channel. Equation (2.5) shows the effect of V_{TH} mismatch in the first channel (Δ_1) and second channel (Δ_{2A}, Δ_{2B}) differential pairs:

$$\begin{bmatrix} i_{aa} \\ i_{ab} \\ i_{ba} \\ i_{bb} \end{bmatrix} = \begin{bmatrix} (V_1+\Delta_1) \\ (V_1+\Delta_1) \\ -(V_1+\Delta_1) \\ -(V_1+\Delta_1) \end{bmatrix} \cdot \frac{g_{m1}}{4} + \begin{bmatrix} (V_2+\Delta_{2A}) \\ -(V_2+\Delta_{2A}) \\ (V_2+\Delta_{2B}) \\ -(V_2+\Delta_{2B}) \end{bmatrix} \cdot \frac{g_{m2}}{4} + \begin{bmatrix} (V_1+\Delta_1)(V_2+\Delta_{2A}) \\ -(V_1+\Delta_1)(V_2+\Delta_{2A}) \\ -(V_1+\Delta_1)(V_2+\Delta_{2B}) \\ (V_1+\Delta_1)(V_2+\Delta_{2B}) \end{bmatrix} \cdot \frac{g_{m1}}{4\eta V_T} \tag{2.6}$$

once this is passed through the recombination stage (effectively multiplied by $\boldsymbol{A^T}$) the result is:

$$\vec{V}_{out} = G_m R_o \cdot \begin{bmatrix} (V_1+\Delta_1) \\ (V_2+\frac{\Delta_{2A}+\Delta_{2B}}{2}) \end{bmatrix} + \frac{G_m R_o}{\eta V_T} \begin{bmatrix} (0) \\ (\Delta_{2A}-\Delta_{2B})\cdot(V_1+\Delta_1) \end{bmatrix} \tag{2.7}$$

Note that only V_{TH} mismatch is considered because the devices are biased in subthreshold. The mismatch in the first stage only causes offsets in the output but not crosstalk from V_2. Crosstalk from V_1 onto the second channel is generated by the mismatch between the differential pairs of the second channel ($\Delta_{2A} - \Delta_{2B}$). Generally, the output current for channel two is $I_2 = g_{m2}V_2 + \alpha_{12}g_{m1}V_1$, where α_{12} represents the crosstalk coefficient from channel one onto

two. α is a Gaussian random variable with zero mean and standard deviation $\sigma_\alpha = \sigma_{VTH}/\eta V_T$, where σ_{VTH} is the V_{TH} mismatch associated with the equivalent parallel input differential pair. For a stack with two or more channels ($n > 2$), each channel can receive crosstalk from upstream channels (but not downstream) with crosstalk coefficients $\alpha_{xi}, x < i \leq n$ all having the same standard deviation, σ_α.

Mismatch in the recombination mirrors also causes crosstalk between the channels. This mismatch alters the weights of Equation (2.3) causing imperfect cancellation of unwanted signals. Unlike crosstalk in the input stack where upstream channels are immune to downstream channels, the recombination mirrors randomly affect all channels. For example, for $n = 2$, the expected crosstalk due to mismatch in subthreshold recombination mirrors on channel two can be modeled as $I_2 = \beta_{12} g_{m1} V_1$. In general, β_{xj} is a Gaussian random variable with zero mean and standard deviation $\sigma_\beta = \sigma_{VTH}/(\eta V_T \cdot \sqrt{2^n})$, where σ_{VTH} is the V_{TH} mismatch associated with a single recombination current mirror device. Thus, for a n-stage stacked amplifier where every channel has an equivalent base g_m, the output voltage signal and crosstalk due to mismatch in the stack and recombination mirrors for a given channel x can be expressed as Equation (2.8). Derivations for σ_α and σ_β is in the appendix.

$$V_{outx} = G_m R_o (V_x + \sum_{i=1}^{x-1} \alpha_{xi} V_i + \sum_{j=1}^{n} \beta_{xj} V_j) \tag{2.8}$$

2.3 System Design

2.3.1 Input Stack

The implemented architecture consists of an input stack, current mirrors, and cascoded output loads (Figure 2.6). Orthogonal current-reuse can utilize any number of stages (n) and stacking more stages will result in a greater power savings, though the number of output currents grows exponentially (2^n). Figure 2.7 shows the implemented orthogonal current-reuse design with four channels (n = 4), resulting in 16 small signal output currents. To match the transconductance between channels, each input PFET device is biased in subthreshold with the same V_{GS} and V_{DS}. To further ensure that each input device in the stack operated at the same bias, the transistor n-wells are tied to their sources to mitigate body effect. The effective width of every channel is the same (192μm/2μm) and is built with an identical differential pair subunits such that the current density of every input device is the same. This matches g_m, g_{DS}, and input capacitance between channels for similar performance. The difference between the DC common mode input levels of the stack also sets the V_{DS} across the input devices. For this design, we chose V_{DS} = 150mV such that $e^{\frac{-V_{DS}}{V_T}}$ affects the drain current by less than 1%. This ensures that the input device bias points remain relatively insensitive to V_{DS} variation from mismatch or input transients. The inputs are DC biased with either external feedback or with on-chip feedback as discussed in Section III-C.

The remaining 900mV (out of a supply voltage of 1.5V) of headroom is consumed by the M_{bp} current mirror and the M_{bn} mirror transistors (Figure 2.7). To save power, the current scaling between M_{bp} and its reference is 16:1. M_{bp} is

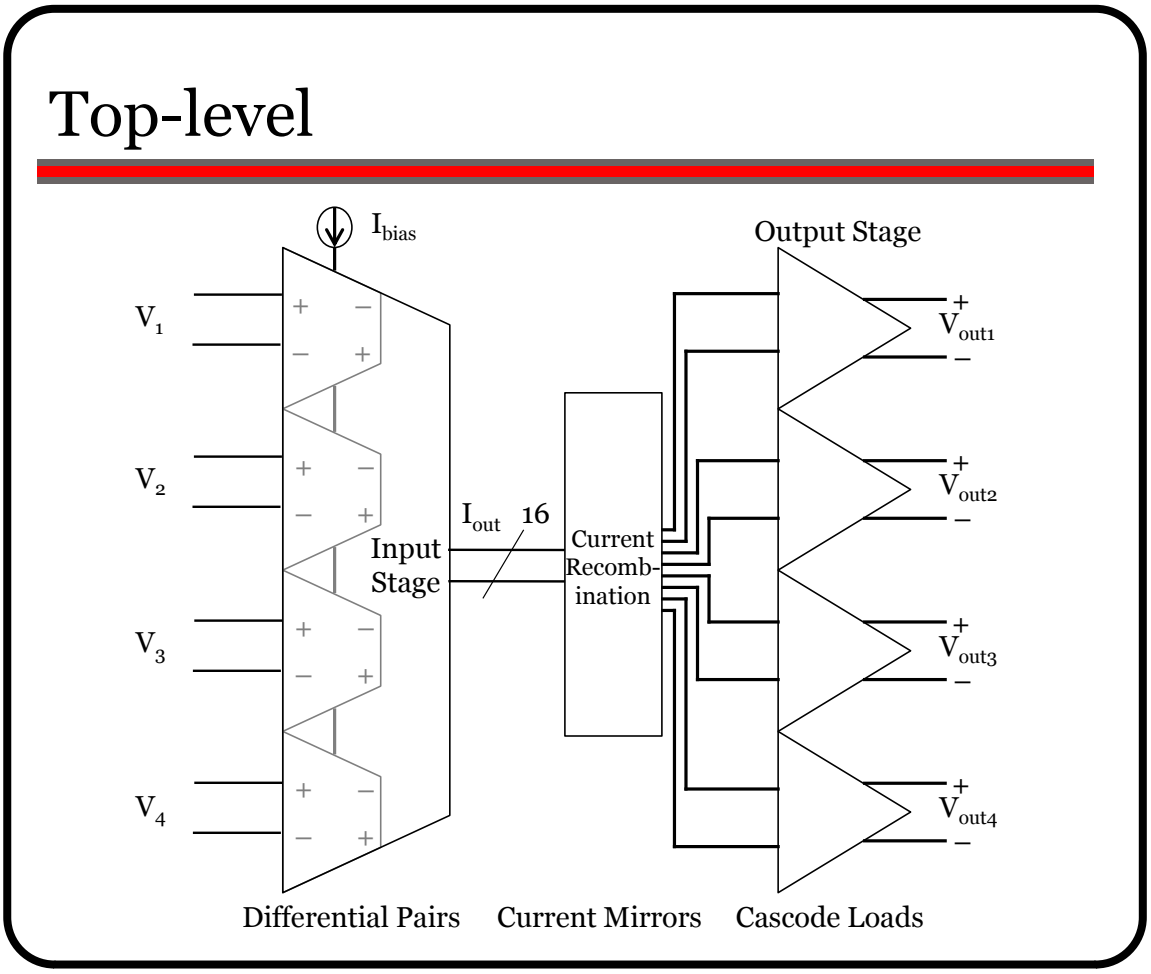

Figure 2.6: Top level diagram of the stacked architecture

cascoded to improve its output impedance for accurate current matching. The M_{bn} transistors have a small W/L ratio to improve current matching without cascoding and to lower their g_m, reducing their noise contributions.

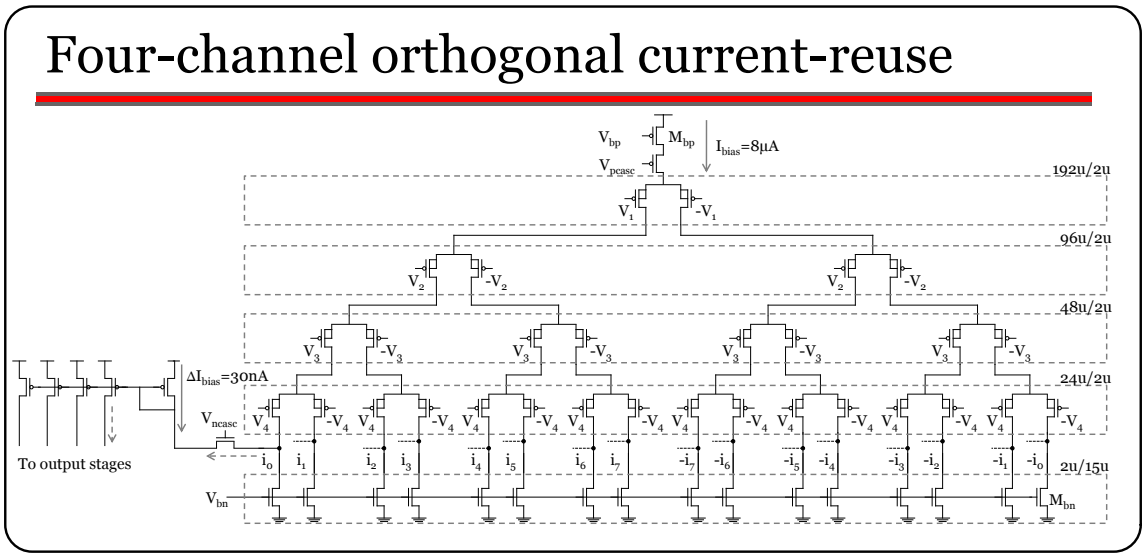

Figure 2.7: Implemented four channel input stack

2.3.2 Signal Recombination

Every output current of the stack is a unique, linear combination of the input voltages. To extract the output currents of the stack, our design uses 16 folded cascode branches to mirror each output current (Figure 2.7). The bias current of each folded branch (ΔI_{bias}) is roughly $1/256^{th}$ of the main stack current to save power and minimize the noise contribution of the folded branch. The minimum ΔI_{bias} current is determined by matching and dynamic range. Since the stack output current due to a differential input is $V_{in}G_m/16$ and inputs can constructively interfere, ΔI_{bias} must be greater than $|\sum_{i=1}^{4} V_{i,in}|G_m/16$. Furthermore, an extreme current scaling ratio between ΔI_{bias} and I_{bias} boosts the input impedance of the folded cascode, attenuating the signal from the bottom differential pair of the stack. Every output current is mirrored four times because each output current is required to construct every output voltage. For example, differential

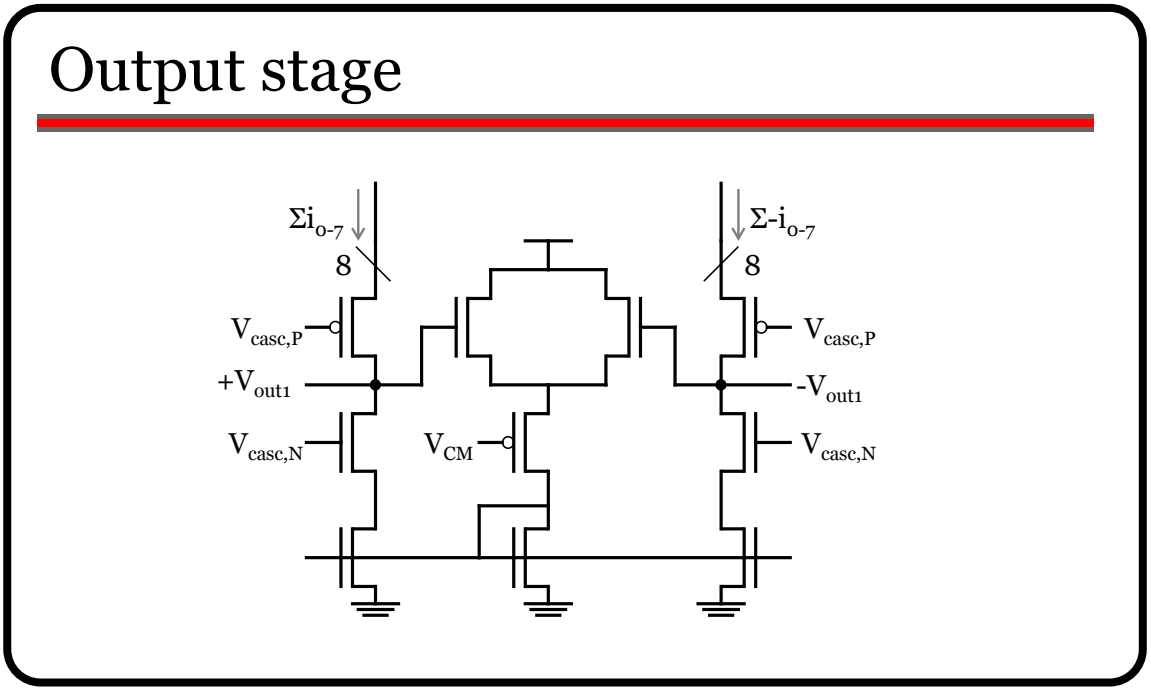

Figure 2.8: Output cascoded load used to recombine signals

output voltage V_{1out} is generated by summing currents i_{0-7} to construct $+V_1$ and currents $-i_{7-0}$ for $-V_1$. Summing currents i_{0-7} cancels the input V_2 components since currents i_{0-3} encode $+V_2$ and currents i_{4-7} encode $-V_2$, each with equal magnitude. Similarly, the V_3 and V_4 components will cancel, allowing V_{1out} to be independent of all inputs but V_1.

Figure 2.8 shows a single differential output stage of the amplifier used to combine the small signal output currents. The schematic also shows the common-mode feedback circuit used to control the output level of the cascoded loads. It uses source followers to avoid resistive loading and is biased at a low current to reduce power consumption. V_{CM} adjusts the DC output level of the cascoded load by controlling the comparison voltage of the common-mode feedback loop. Each output common mode is separately controlled to allow

separate control of input common mode when biased through feedback.

2.3.3 On-Chip Feedback

Each amplifier in the stack is fully-differential and has high open-loop gain (85dB, simulated). Since every channel is independent, our design has the option to implement on-chip capacitive feedback using MIM capacitors as shown for a single channel in Figure 2.9. The inputs are AC-coupled to set the DC bias level of each input independently. Pseudoresistors [18] with high incremental resistance (r_{inc}>100GΩ) are used to provide DC feedback to set the input level equal to the output level. Additionally, the pseudoresistors set a low-frequency high-pass corner ($\omega_H = 1/r_{inc}C_{in}$) to block electrochemical offsets but not local field potentials.

2.4 Measurement Results

Our amplifier was fabricated in a 0.13μm process through MOSIS. The circuitry, including bias circuitry and amplifiers, occupies an area of 0.0264mm^2. Including all 16 feedback capacitors, the occupied area increases to 0.125mm^2, which corresponds to .03125mm^2 per amplifier. The on-chip feedback was designed for a gain of 100, with C_{in} = 20pF and C_f = 200fF with a density of 2fF/μm^2. A die photo is shown in Figure 2.10.

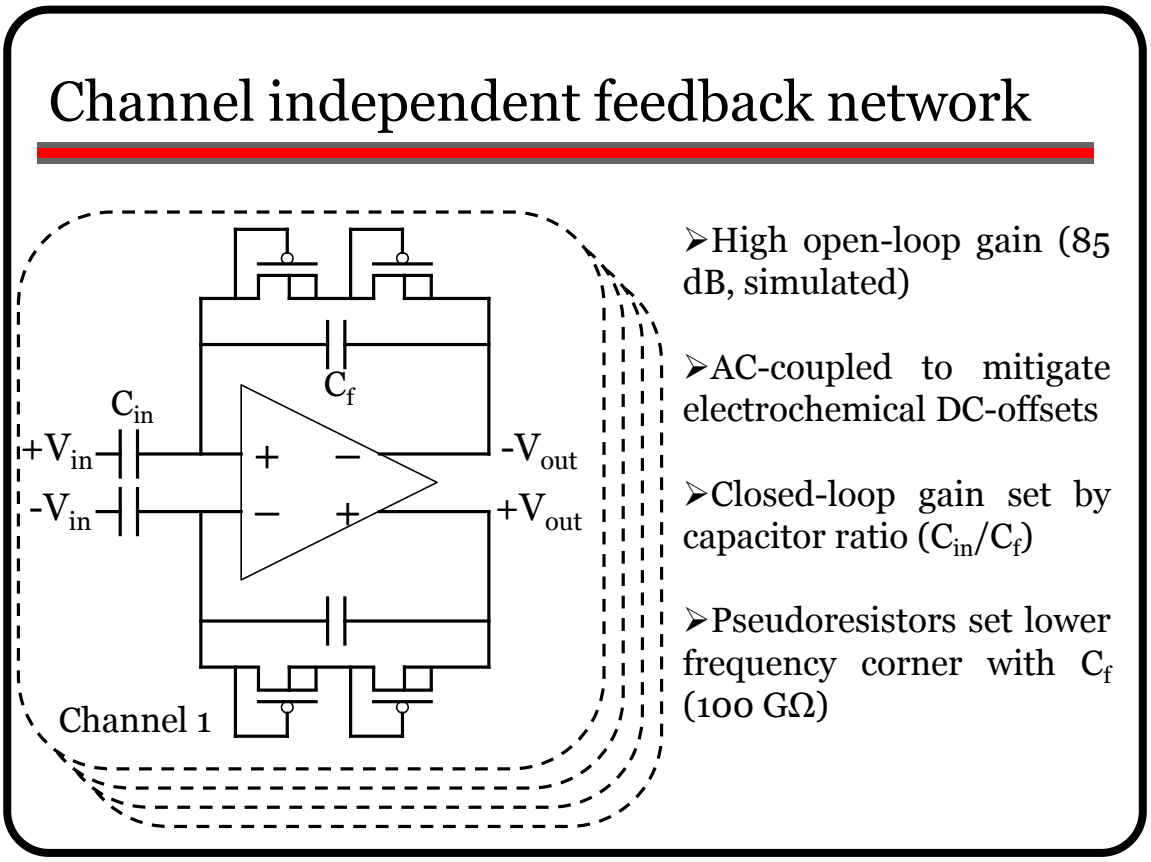

Figure 2.9: Feedback network for each channel

2.4.1 Transfer Functions

Figure 2.11 shows the measured transfer functions of all four amplifiers in feedback. The bandwidths range from 14.1kHz to 19.9kHz with I_{bias} = 8μA and ΔI_{bias} = 30nA. The mid-band gain is approximately 40dB, with less than 1dB variation between channels. The bandwidth variation between the channels is likely due to several factors. The impedance seen by the fourth channel's differential pairs is the $1/g_m$ of the folded cascode transistor, which is much higher than the impedance seen by the other channels, reducing its effective G_m. Conversely, signals at the top of the stack are attenuated as they pass through more cascode

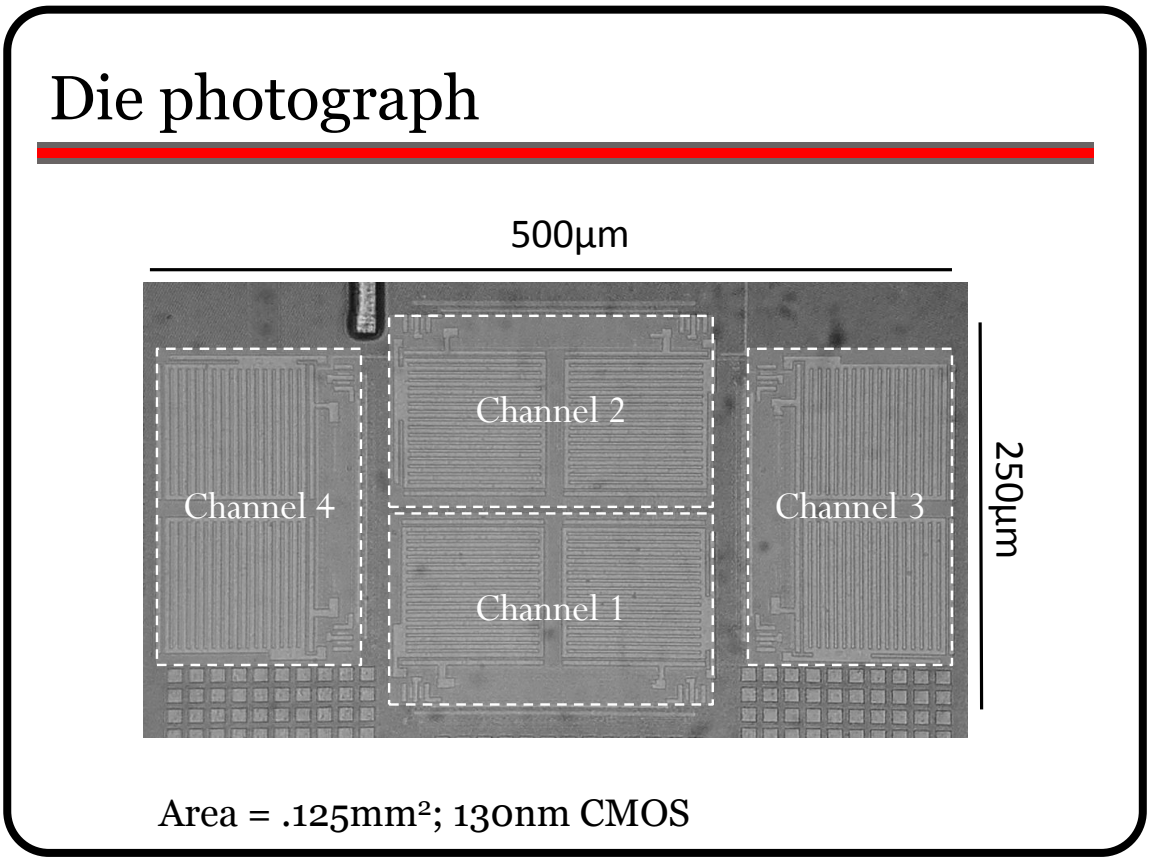

Figure 2.10: Stacked amplifier die photograph

stages. Furthermore, since the different output stages are biased at different common mode voltages, some variation in gain and BW is to be expected.

2.4.2 Input-Referred Noise

Figure 2.12 shows the input-referred noise spectra of all four amplifiers in feedback. The flicker noise corner is 1kHz and the thermal noise level is 25nV/ $\sqrt{Hz}$. Integrating the noise power from 10Hz to 100kHz results in an input-referred noise ranging from 3.4μV_{rms} to 4.2μV_{rms} across channels. Since the input stack

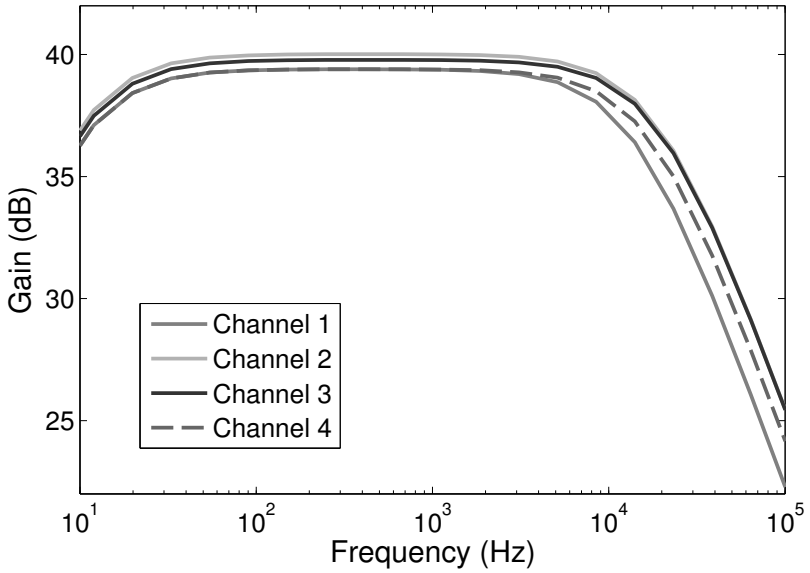

Figure 2.11: Frequency response of all four amplifiers in the stack.

consumes 8μA, the cascoded mirrors consume 490nA, and each output stage consumes an additional 490nA, the NEF of a single pathway ranges from 3.0 to 3.7, similar to other differential amplifiers reported [87], [93]. Shorting the inputs of all four channels and averaging the outputs in the time domain results in an input-referred noise of 1.9μV_{rms}. This results in a NEF of 1.75, the lowest reported to our knowledge, even for single-ended topologies (Table 2.1). This result can also be seen by using an effective NEF [42], calculated for each channel by amortizing the shared bias current by dividing it by the number of channels. This comparison is used in Table 2.1 to calculate effective NEF ranging from 1.64 to 1.97 across channels. A modified power efficiency metric PEF = ($NEF^2 \cdot VDD$) defined in [67] can also be used to account for noise, current and required voltage headroom. The PEF of a single channel was 13.9 and calculated to be 4.0 using effective NEF, demonstrating significant enhancement compared to previously reported differential amplifiers and competitive with

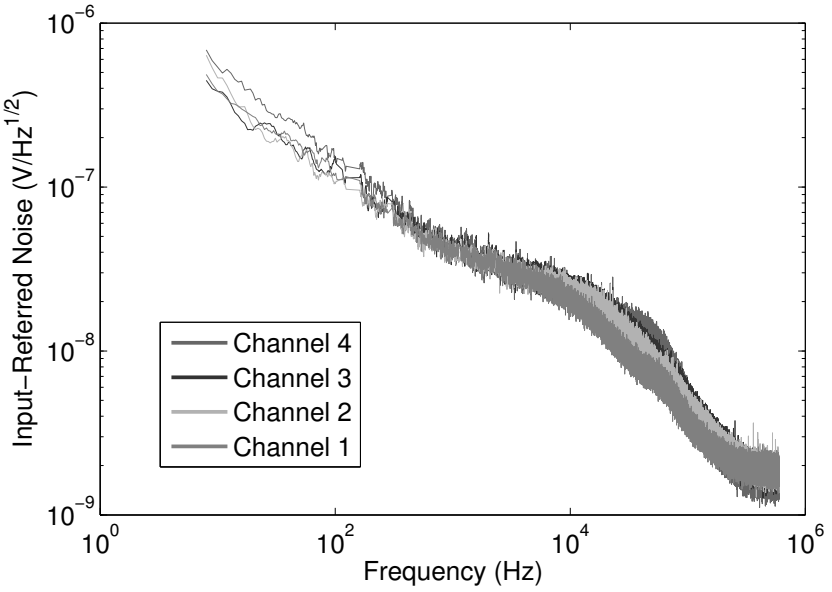

Figure 2.12: Input-referred noise of all four amplifiers.

single-ended amplifiers.

2.4.3 Amplifier Nonlinearity and Crosstalk

The total harmonic distortion (THD) of a single channel on our amplifier was 1% at an input level of 16.5mVp-p, which is sufficient to accommodate the largest extracellular biopotentials (<1mVp-p) with significant margin for error. The measured CMRR and PSRR were 78dB and 80dB, respectively. Therefore, the concern is not the performance of a single channel, but how the channels interfere with one another. The proximity and interconnection of the four stacked amplifiers increases the risk of crosstalk between the amplifiers. Ideally the output voltages are entirely independent, with each output responding to only one input. In reality, device mismatch and any other imbalances will cause some crosstalk between pathways. Linear crosstalk for the stack was measured by

Table 2.1: Comparison of Power Efficient Amplifiers

	JSSC '12 [67]	BioCAS '12 [93]	EMBC '07 [41]	BioCAS '07 [87]	This work (Single Channel)	This work (Effective Channel)
Power	4.13μW	12.1μW	0.8μW	7.56μW	13.4μW	3.9μW
VDD	0.5V	1.0V	1.0V	2.8V	1.5V	1.5V
Noise	4.9μV_{rms}	2.2μV_{rms}	3.6μV_{rms}	3.1μV_{rms}	3.7μV_{rms}	3.7μV_{rms}
Bandwidth	10kHz	10.5kHz	4.7kHz	5.3kHz	19.9kHz	19.9kHz
Gain	32dB	40dB	36.1dB	40.85dB	40dB	40dB
THD	2% @200μV_{rms}	1% @1mV_{p-p}	7.1% @1mV_{p-p}	1% @7.3mV_{p-p}	1% @16.5mV_{p-p}	1% @16.5mV_{p-p}
NEF	5.3	2.9	1.8	2.67	3.04	1.64
PEF	14.05	8.4	3.24	19.96	13.86	4.03
CMRR	75dB	80dB	–	66dB	78dB	78dB
PSRR	64dB	80dB	5.5dB	75dB	80dB	80dB
Technology	65nm	130nm	500nm	500nm	130nm	130nm
Topology	Fully-Differential	Fully-Differential	Single-Ended	Differential	Fully-Differential	Fully-Differential
Area	.0037mm^2	.072mm^2	.046mm^2	.16mm^2	.03125mm^2	.125mm^2

injecting a 5mVp-p 3.5kHz sine wave into one input at a time and measuring the 3.5kHz signal power of the other amplifier outputs. Table 2.2 shows the measured crosstalk for one chip where the diagonal terms represent the desired gain terms in dB and the off-diagonal terms correspond to the crosstalk at the fundamental frequency. We measured crosstalk on all channels of two samples and found no systematic difference between channels. Crosstalk coefficients appeared to be roughly gaussian in distribution, with a standard deviation of 0.39 input-referred (Figure 2.13). Assuming a threshold mismatch of 2.2mV-μm for our process [69], our predicted σ_α was 0.27 and σ_β was 0.39 based on device sizing. This implies that the mismatch of the recombination mirrors should dominate crosstalk, with an RMS value consistent with our measurements.

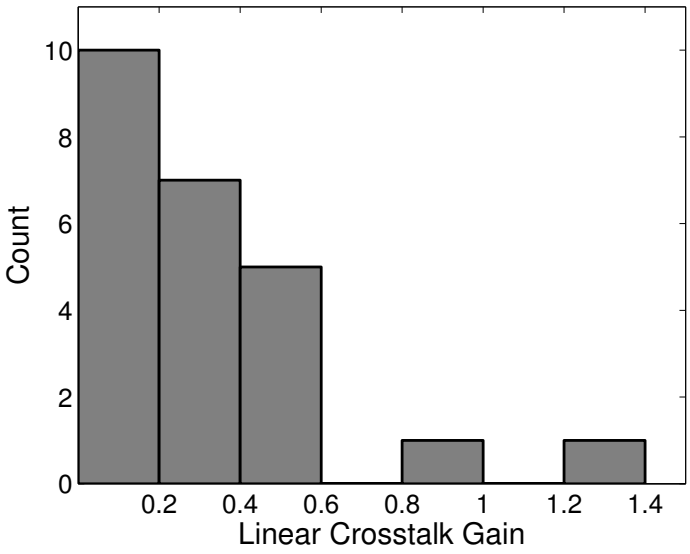

Figure 2.13: Histogram of linear crosstalk gain between channels

Table 2.2: Gain and Fundamental Frequency Crosstalk Between Amplifier Channels

V_{in}@5mV$_{p-p}$	V_1	V_2	V_3	V_4
V_1out(dB)	39.4	-8	-32	-10
V_2out(dB)	-35	40	-12	-20
V_3out(dB)	-9	-10	39.8	-11
V_4out(dB)	-7	-7	-9	39.4

At 5mVp-p, the distortion is dominated by soft-nonlinearity rather than saturation of the folded cascode. Inter-channel nonlinearity was also characterized by measuring the intermodulation between channels. Intermodulation was tested by injecting a 5mVp-p 3.5kHz sine wave into channel one and a 5mVp-p 2kHz sine wave into the other channels. Table 2.3 shows the measured intermodulation between the channels relative to the 2kHz output signal, with all intermodulation terms small enough to be negligible.

Table 2.3: Intermodulation Between Channels

V_1@5mV$_{p-p}$	V_2	V_3	V_4
3.5 kHz fund.(dB)	-75	-49	-47
5.5 kHz IM2(dB)	-75	-69	-67
9 kHz IM3(dB)	-72	-76	-71

2.4.4 Neural Data

To confirm that our amplifier could operate as four independent channels that could provide robust, simultaneous low noise recordings, we used it to record endogenous spiking activity from a 300μm-thick mouse olfactory bulb slice. Recordings were taken with a planar microelectrode array (Multi Channel Systems MCS GmbH) with 60 TiN electrodes. The electrodes have a 30μm diameter, are spaced by 200μm and have an impedance of 50kΩ at 1kHz. A subset of four non-adjacent electrodes were chosen to avoid recording highly correlated spiking activity. A long recording and a short recording of endogenous activity from all channels are shown in Figure 2.14.

2.5 Metrics for Power-Efficient Amplifiers

The proposed approach clearly provides a significant benefit in terms of NEF. However, simply stacking amplifiers would also provide a similar benefit by simply trading bias current for voltage headroom without any actual benefit in terms of power efficiency. In order to provide a more useful measure of the power efficiency of this approach, we propose two alternate measures of noise-power efficiency, similar to the power efficiency factor (PEF) defined in [67].

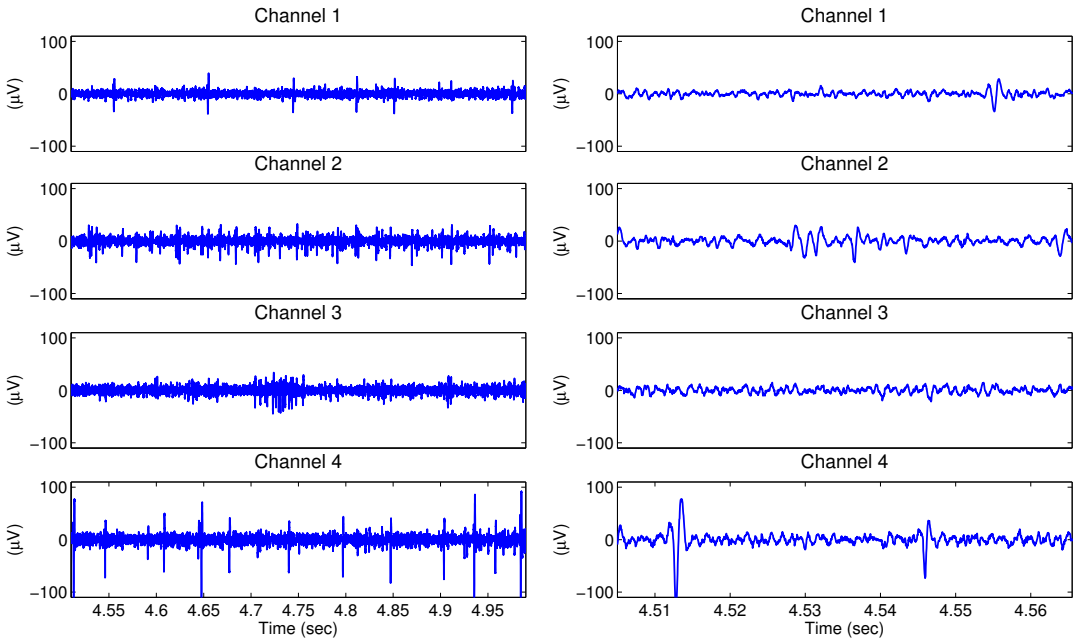

Figure 2.14: Neural recording from a mouse olfactory bulb.

2.5.1 Noise-gain Efficiency Factor

Following on the idea of the NEF, we can define a noise-gain efficiency factor, NGEF, defined as the DC power consumed by an amplifier relative to the power consumed by an ideal subthreshold MOSFET ($\gamma = 1$) in common-source configuration with identical noise and voltage gain. Similar to derivations of NEF, we can calculate the ideal current based from the measured input-referred noise and bandwidth:

$$I_{ideal} = \frac{4kT \cdot V_T \cdot BW}{V_{rms,in}^2} \cdot \frac{\pi}{2}. \tag{2.9}$$

To find the minimum possible voltage bias, we note that while the noise of a transistor is not dependent upon V_{DS}, the voltage gain is. In subthreshold at

small V_{DS}, it can be shown that

$$g_{ds} = \frac{I_D}{V_T} \cdot e^{\frac{-V_{DS}}{V_T}} \tag{2.10}$$

and maximum voltage gain, $A_v = g_m / g_{ds}$. We can therefore calculate the ideal minimum voltage from the measured amplifier voltage gain:

$$V_{ideal} = V_T \cdot \ln |A_v| \,. \tag{2.11}$$

Thus, for a given input voltage noise and gain, the overall power efficiency is given by the measured power divided by the product of the ideal voltage and current:

$$NGEF = \frac{I_{measured} \cdot V_{measured}}{I_{ideal} \cdot V_{ideal}} = NEF^2 \cdot \frac{V_{DD}}{V_T \cdot \ln |A_v|}. \tag{2.12}$$

While NEF quantifies how well an amplifier maximizes g_m for a given bias current, NGEF additionally quantifies how well an amplifier maximizes A_v for a given voltage supply.

NGEF = 1 is only attainable by an ideal BJT with a large inductive load. Practically, a CMOS inverter in resistive feedback, self-biased in deep subthreshold, provides a nearly minimum NGEF at low frequencies. While stacking NMOS and PMOS devices uses twice the headroom of a single transistor with the same gain, the transistors' transconductances act in parallel, effectively doubling g_m [66] and so halving the required current for a given noise level. Since each transistor acts as a load for the other transistor, no power is wasted on active loads

or tail current sources. These various benefits help explain the comparatively low values NGEF acheived by single-ended, complimentary designs [41].

Many applications require differential topologies to reject power supply and common-mode noise sources, doubling the NGEF compared to single ended designs [93]. More traditional differential, current-biased designs are intrinsically less efficient than an idealized complimentary design because of the added overhead of tail currents and active loads and folded cascodes [87]. However, orthogonal current stacking can dramatically mitigate these losses by guaranteeing that the majority of the voltage headroom is used by active input transistors rather than biasing transistors, resulting in an NGEF on par with the lowest power single-ended complimentary designs while also maintaining a fully differential design (Table 2.1).

2.5.2 Dynamic Range Efficiency Factor

From a pure small-signal perspective, NGEF provides a good measure of how close a single-stage amplifier comes to ideal power for a given level of performance. However, in real amplifiers, the real limit on power consumption is often set by the instantaneous dynamic range of the amplifier. That is, power is set not only by the minimum signal that must be distinguishable from noise, but also the largest signal that can be handled simultaneously. This suggests an alternate measure of efficiency, based on noise and dynamic range.

In a high-gain amplifier, such as used in neural recording, even small input voltages can generate large outputs that could saturate the amplifier. In this case, again for a simple common-emitter amplifier, the maximum output

voltage signal is $V_{omax} = V_{DD}/2$, reflecting a rail-to-rail swing. Thus, for a given noise floor and output swing, one can define an ideal power consumption $P_{ideal} = 2 \cdot I_{ideal} \cdot V_{omax}$ corresponding to a dynamic-range efficiency factor. If we define V_{omax} to be that signal strength that generates 1% total harmonic distortion (THD), then we see that fully-differential designs gain a significant advantage, both by suppressing even-order distortion, and by providing effectively twice the (differential) voltage swing for the same supply voltage.

$$DREF = \frac{I_{measured} \cdot V_{DD}}{I_{ideal} \cdot V_{omax}} = NEF^2 \cdot \frac{V_{DD}}{2 \cdot V_{omax}} \tag{2.13}$$

Under this figure of merit, orthogonal stacking provides a dramatic benefit, more than 10x that of other designs (Table 2.1), allowing for very efficient differential circuits while by amortizing the primary bias current across four channels but still maintaining full, differential output swing.

2.6 Other Implementations

This section describes additional enhancements, alternate implementations, and applications of the orthogonal current reuse amplifier topology.

2.6.1 Current Chopping

Voltage chopping at the input and output of the amplifier imposes two constraints. Firstly, the chopping period should be much greater than the settling time of the amplifier to avoid decreasing the open-loop gain (A_{eff}=A_o(1-4τ/T)) [19]. This is problematic in energy efficient amplifiers where the bandwidth

may be close to the $1/f$ noise corner and the chopping frequency (f_{chop}) must to be higher than the noise corner to be effective. Thus, the additional power required to increase the amplifier bandwidth makes chopping difficult for energy efficient designs. Secondly, the input impedance is drastically lowered by a switching input with a large input device capacitance and high f_{chop}. Decreasing the input device size increases the intrinsic $1/f$ noise and offset while decreasing f_{chop} may bring it too close to the $1/f$ noise corner. Additionally, decreasing f_{chop} increases ripple and requires multi-path architectures to suppress ripple [23]. To circumvent these issues, we implemented internal current chopping. Due to the differential topology of the input stack, every small signal output current has a compliment. This enables internal chopping modulation to be implemented via switches at the output of the input stack as seen in Figure 2.15, and at the output cascode stage. In Figure 2.15, current from terminal A_0 is steered to i_0 or -i_0 through transmission gates with 50% duty cycle squarewave control. Chopping the internal nodes steers the complimentary currents allowing for faster modulation beyond the amplifier closed-loop bandwidth. This means that the modulation frequency is useful for the spike frequency as well as LFP bands. The output pole low pass filters the up-modulated intrinsic amplifier offsets and $1/f$ noise.

2.6.2 Wireless Rectifier

The ultimate goal of low-power, low-noise amplifiers for neural recording is to implement them in wireless systems. For this reason, many power-efficient amplifiers are trending toward architectures with lower supply voltages (VDD < 0.5V) to be more amenable to wireless power transfer. However, the orthogonal

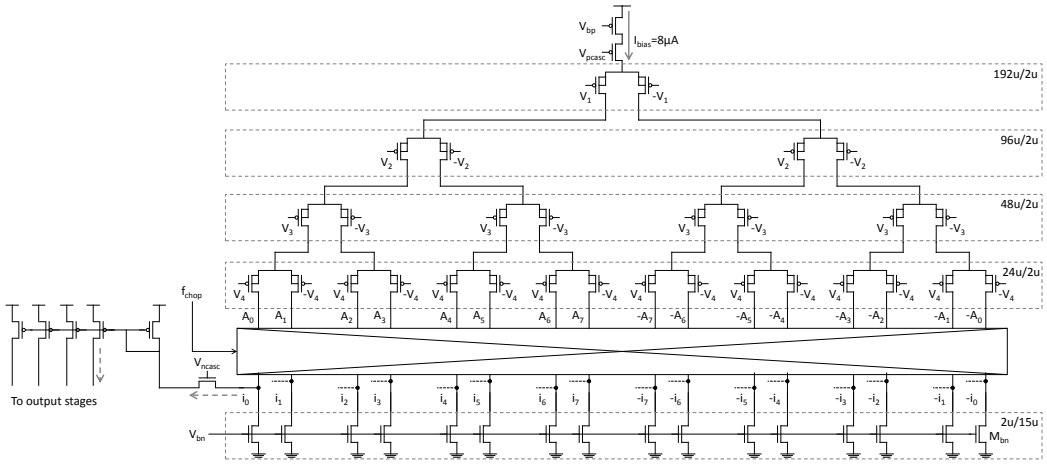

Figure 2.15: Input stack with output current chopping

current reuse amplifier requires increasing the supply voltage to reap benefits in power efficiency (Fig. 2.16). In [54], we presented a wireless, low-voltage ladder rectifier that utilizes the advantages of the orthogonal current reuse amplifier topology [44].

Inductively coupled power transfer has been widely used [52, 91, 80], especially at low frequencies (< 20MHz) where the tissue absorption of field energy is negligible [70]. Nonetheless, previous systems have lacked the ability to convert low voltage inputs (<200mV) on the inductor. Furthermore, for bio-medical monitoring device applications, it is desirable to minimize the overall device size, which, for inductively powered devices, is limited by the size of coil used. However, input power to the implants rectifier is limited by the size of the inductor, creating a trade-off between available power and the device dimensions. Moreover, while extremely low power circuits have been demonstrated with reduced inductor size, such circuits still require hundreds of millivolts of DC voltage to function, and typical rectifiers themselves require even greater AC

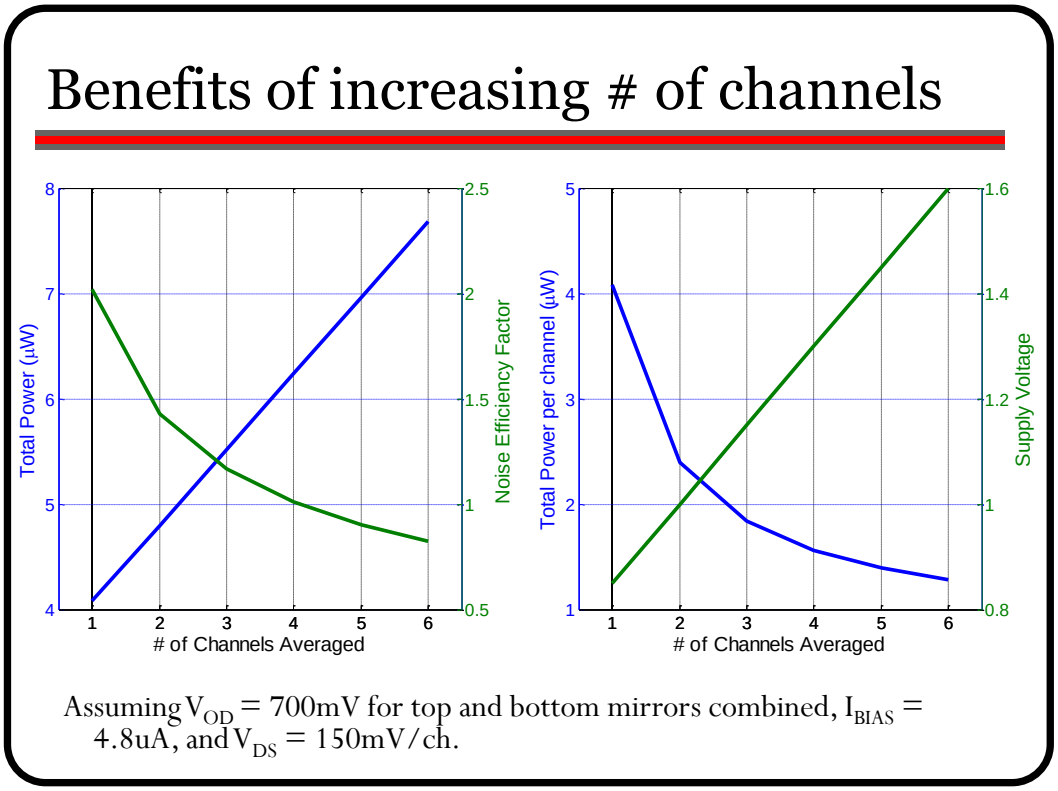

Figure 2.16: Input stack with output current chopping

voltages to be able to efficiently generate the needed DC voltages. Since output voltage is also a function of coil inductance (and so, coil size), there is a need for implantable power systems which can make use of low-voltage, low-power AC signals. In [54], we demonstrated a novel rectifier consisting of a full-wave rectifier, and cascaded self synchronous charge-pump stages to provide high voltage DC output from a small inductor which can power an energy efficient current-reuse amplifier for neural recording (Figure 2.17).

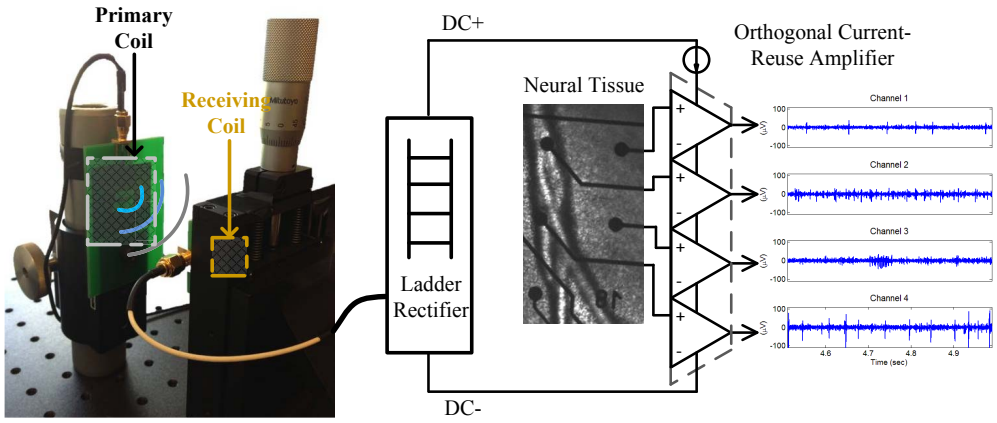

Figure 2.17: Wireless power transfer configuration

2.6.3 Low-Power Receiver

Passive-mixer first receivers have become ubiquitous with software-defined radios due to their inherent high wideband linearity, low noise, and wide tuning range. In [5], we presented a high performance, low power, passive-mixer first receiver with resonant LO generation and an orthogonal current reuse baseband amplifier with built-in odd harmonic rejection (Figure 2.18).

Passive-mixer first receivers are favorable for low power design since very little circuitry operates at RF bands, making use of slower, more efficient circuits. Furthermore, multi-phase radios require multiple signal pathways (4 differential signals for 8-phase), meaning they can utilize a four channel orthogonal current reuse amplifier. In our implemented design, we incorporated harmonic rejection into the baseband amplifier by combining adjacent phases (Figure 2.19). The weighting for the adjacent phases was $1/\sqrt{2}$ relative to the fundamental in order to suppress the 3rd and 5th harmonics.

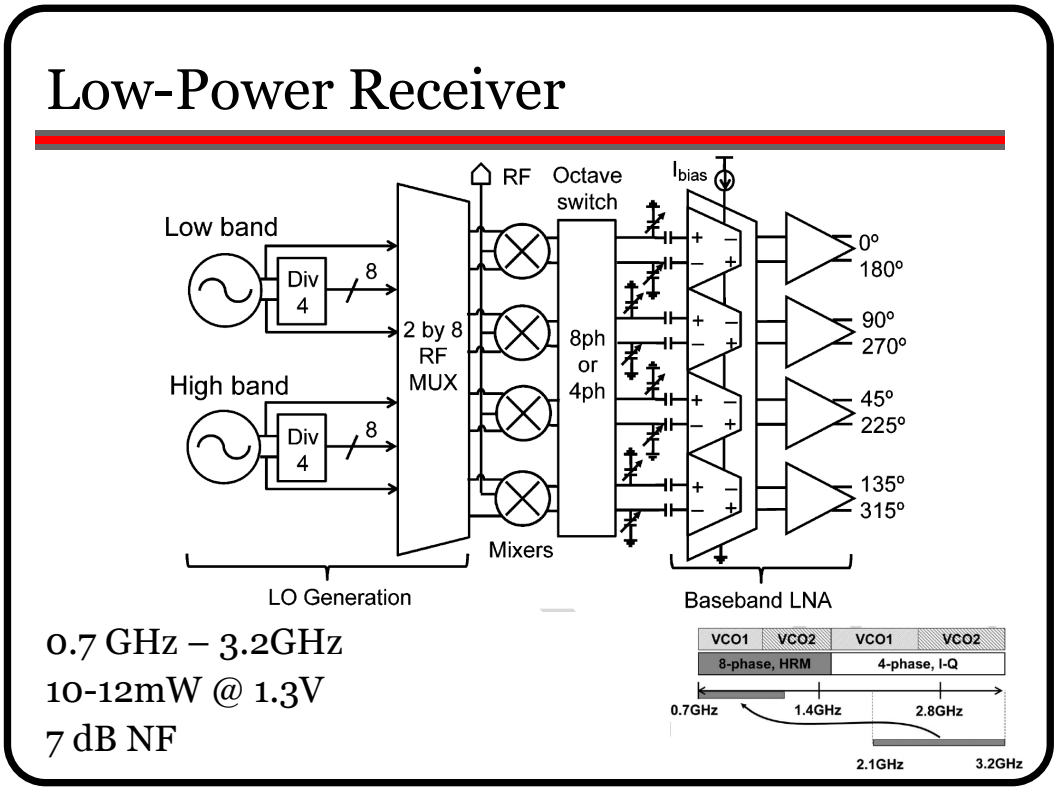

Figure 2.18: Block diagram of the low-power receiver

2.7 Conclusions

This work has presented a new approach to minimizing the noise-power trade off in integrated amplifiers with multiple inputs. By reusing current between four parallel stages, significant power savings was demonstrated without degrading noise, resulting in a better effective noise efficiency factor than previously reported. Compared to simply lowering the supply voltage of the input pair on 4 separate amplifiers, this approach provides significant savings by amortizing the bias voltage overhead of the various DC current sources in the amplifier. Additionally, our approach saves power in the input stages and accommodates a large output swing with a single supply voltage. To suppress

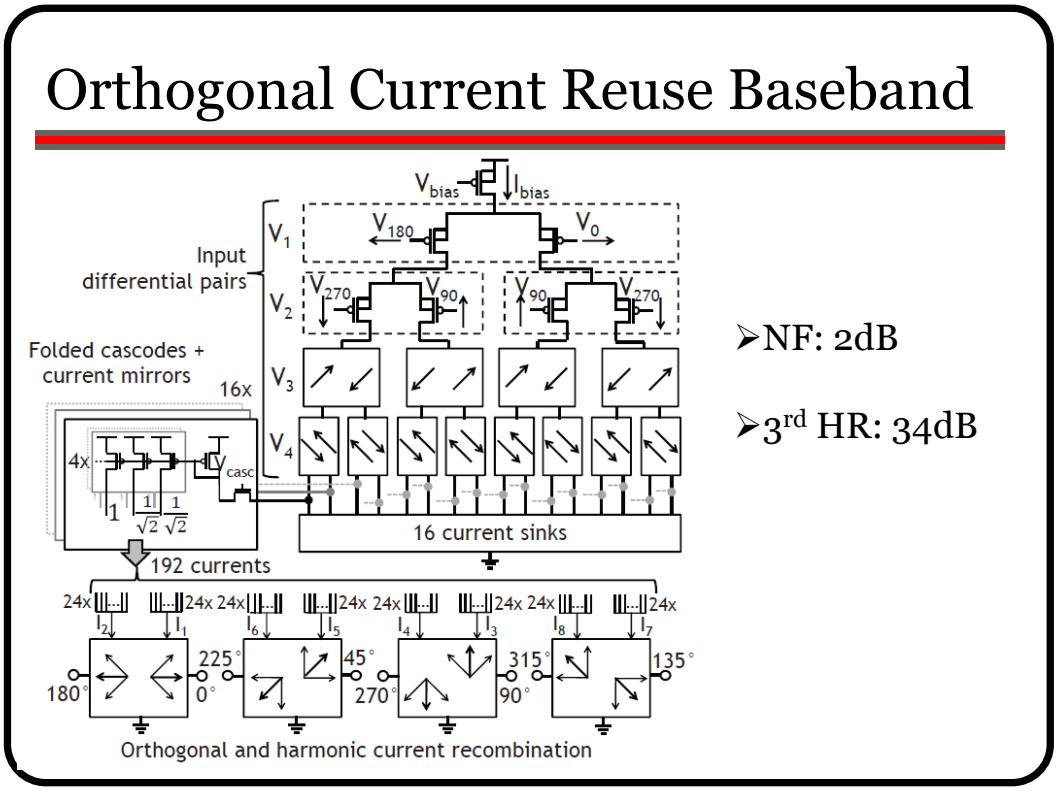

Figure 2.19: Schematic of the baseband amplifier

crosstalk between stacked amplifiers, current was distributed between layers in an orthogonal fashion, isolating the interleaved amplifiers by better than 40dB under 3-σ mismatch. Due to noise from the electrode, bulk electrolyte, and neuronal noise, the dynamic range of extracellular neural signals typically range from 24dB to 42dB [59]. Thus, the noise performance and crosstalk isolation of our amplifier are sufficient for neural recordings. Higher performance applications, however, may require better isolation between channels, requiring better matching of components, or calibration in post-processing. The resulting output currents are combined in output stages that are not stacked so that they are capable of large output swings, but also consume little current and contribute little overall noise. A stacked amplifier is sensible for any application where

multiple low frequency signals need to be amplified with low noise at minimum power. This application space ranges from parallel sensing applications [42] to multiphase radio baseband amplifiers [4].

CHAPTER 3

BIOLOGICAL PROCESSING: OSCILLATORY DYNAMICS OF OLFACTORY BULB SLICES ON A PLANAR MULTIELECTRODE ARRAY

3.1 Introduction

Synchronized activity across neural assemblies produces oscillations in the local field potential (LFP) indicative of biophysical constraint in the temporal patterning of coactivated neural populations. While little evidence suggests oscillating field potentials significantly influence underlying cellular activity (Anastassiou et al., 2010 [3]) there is growing evidence that disruption to the circuitry responsible for temporal coordination underlying oscillating field potentials can produce deleterious effects in the propagation of information as observed in insect olfactory representations (Stopfer et al., 1997 [79]; MacLeod et al., 1998 [63]), spatial memory learning (Girardeau et al., 2009 [33]), and grid cell functionality (Brandon et al., 2011 [6]). More specifically, oscillations in the 30-120Hz range, termed gamma oscillations, are a feature of many stimulated vertebrate brain regions, particularly the olfactory bulb (OB; (Adrian, 1950 [1]; Kay and Freeman, 1998 [51]), occipital lobe (Gray et al., 1989 [35]), and hippocampus (Traub et al., 1996 [82])). Hence, the induction of LFP oscillations can provide insights into the dynamics of coordinated ensembles of neurons and help identify the biophysical properties that underlie the precise temporal regulation of widespread population activity.

Communication across multiple, costimulated brain regions has been hypothesized to occur through phase-locking of simultaneously but segregated oscillatory networks resulting in coherent field oscillations (Fries, 2005 [29]).

Coherent gamma oscillations in the rodent OB are a conspicuous trademark of OB recordings (Freeman, 1978 [25]; Kay and Freeman, 1998 [51]). Coherent bulbar activity is produced by odor investigation, during which, the OB generates odor representations based on the afferent activation patterns of primary olfactory sensory neurons (OSNs). These axons converge onto spatially discrete glomeruli within the OB that each represent the activation of a single class of odorant receptor (Buck and Axel, 1991 [12]). The OB transforms these primary afferent representations through decorrelation across similarly activated glomeruli and normalizing input activation patterns to levels within the dynamic range of central neural circuitry (for review see; (Cleland, 2010 [14]). This functions to create a secondary odorant representation hypothesized to be dependent upon a precise temporal regulation in the action potential generation of the OBs primary principle output layer, mitral cells (Linster and Cleland, 2010 [58]; Gschwend et al., 2012 [36]), yielding coherent, 40-100 Hz gamma band oscillations throughout the OB (Kay and Freeman, 1998 [51]; Kashiwadani et al., 1999 [50]; Beshel et al., 2007 [9]). Coherent LFP oscillations across the OB indicates the presence of a global, temporally precise mechanism capable of synchronizing spatially segregated but similarly activated glomerular activity; thus, making the olfactory bulb an advantageous neural system to study synchronized circuit dynamics.

To quantify the extent of phase-locked gamma oscillations across a reduced preparation of the rodent OB, we use a 60-channel planar, microelectrode array (MultiChannel Systems) to record spatially distinct, long-lasting gamma oscillations induced by pharmacological methods shown to induce gamma oscillations in the hippocampus (for review see; (Whittington et al., 2000 [88]). In addition to pharmacological methods, we used tetanic optical stimulus on slices

from OMP/ChR transgenic mice to induce gamma oscillations. We found the majority of gamma oscillations in the OB slice occurred in the 20-55Hz range, consistent with other OB gamma recordings (Friedman and Strowbridge, 2003 [28]; Galan et al., 2006 [30]; Gire and Schoppa, 2008 [34]), and maintained phase-locked synchrony up to 280 μm persisting often for several minutes.

3.2 Materials and Methods

3.2.1 Olfactory Bulb Slices

Horizontal slices were taken from the olfactory bulbs of wildtype CD-1 and OMP/ChR transgenic mice aged 4- to 6 weeks. Mice were rendered unconscious with isoflurane gas before they were anaesthetized with ketamine. Three hundred micron thick slices of olfactory bulbs were taken using a vibrating microtome and incubated in warmed (35°C-38°C), oxygenated artificial cerebral spinal fluid (aCSF) dissecting solution with reduced calcium (add concentrations of ions). After twenty minutes of incubation, OB slices were returned to room temperature for storage until transferred to the recording well on the electrode array chip where they were perfused (1 mL/minute) with heated (34°CC), oxygenated aCSF (add concentrations of ions) from a gravity feed perfusion system. Slices were held in position by nylon webbing glued to a "'C"' shaped chrome wire.

3.2.2 MEA Electrophysiology

Extracellular data were recorded from the OB slice using a 60-electrode planar microelectrode array (MEA; Multi Channel Systems, Germany). The electrode array had sixty titanium nitride electrodes (200μm pitch, 30μm width, 30-50$k\Omega$ impedance) arranged in an 8x8 (minus the four corners) grid with one of the electrodes operating as a reference electrode. The array connected to the Multi Channel Systems baseplate (MEA1060) with each electrode individually connected to an amplifier (1200x gain).The baseplate and electrode array chip connected to a PCI data acquisition system and data were acquired using the MC_RACK program. Signals were bandpass filtered (1Hz-3000Hz), amplified, and sampled (5-20kHz analog-to-digital).

3.2.3 Pharmacology

Slices were stimulated by pharmacological agonists pipetted directly onto the slice in the recording well. In these experiments, we used an acetylcholine receptor agonist (carbachol), a metabotropic glutamate receptor agonist (ACPD), and a group I metabotropic glutamate receptor agonist (DHPG) that were all purchased from Sigma-Aldrich.

3.2.4 Optical Stimulus

Slices were stimulated by 4Hz tetanic bursts of light (475nm) generated by a light engine directed onto the slice for a duration of 5s.

3.2.5 General Analysis

Analyses were performed with MATLAB. Data were resampled at 512Hz. Principal component analysis was performed on array-wide datasets to remove 60Hz interference and other correlated noise sources. Spectrograms were computed with fast Fourier transorms (FFT) using 1sec time intervals convolved with a triangular window. Intervals overlapped by 50% (500msec). Spectrograms were smoothed using a two-dimensional hamming window (16 points, 4Hz by 4sec). LFP gamma band data were bandpass filtered with a 2nd order butterworth filter (20Hz $\leq f_{BP} \leq$ 55Hz). Biased autocorrelations were performed on 250msec long, bandpass filtered data.

3.2.6 Gamma Detection

Isolating persistent gamma oscillations was performed by thresholding spectrograms at 2 standard deviations above the mean power and clustering connected components. Clusters were bridged via 1 standard deviation pixels then sufficiently large clusters (>5sec) were identified as possible gamma oscillations. The clusters were then used as a mask on raw spectrograms for pathfinding. Beginning and end points were manually selected on the spectrogram and the path determined by the highest power pixels within the mask. Oscillations whose frequencies were stable (<5Hz variation) for 80sec were selected for analysis.

3.2.7 Gamma Analysis

The LFP gamma oscillation settling frequency was determined by the mean frequency during the 80sec window. The integrated gamma band power was used to evaluate effect of drug application to slice using an 80sec analysis window before stimulation and an 80sec after stimulation during the detected stable oscillation. The gamma power (P_γ) was calculated by

$$P_\gamma = \int_{20}^{55} X(f)df, \tag{3.1}$$

where $X(f)$ is the FFT of the 80sec analysis window. Statistical significance was calculated using Students t-test on the pre- and post-stimulus integrated powers. Overlaid power spectra were computed with the same analysis windows using Welchs power spectral density estimate and then normalizing the post-stimulus density (P_{post}) at frequency f to the pre (P_{pre}):

$$\Delta(f)\% = \frac{P_{post}(f)}{P_{pre}(f)}, \tag{3.2}$$

yielding $\Delta(f)\%$, which is the percentage change in power density. Coherence (C_{xy}) between oscillations on different electrodes was computed in the frequency domain via

$$C_{xy}(f) = \frac{\left|\langle P_{xy}(f)\rangle_n\right|}{\sqrt{\langle P_{xx}(f)\rangle_n \cdot \langle P_{yy}(f)\rangle_n}}, \tag{3.3}$$

where P_{xx} and P_{yy} are autocorrelations and P_{xy} is the cross-correlation between the two electrodes, each averaged over n, 1sec epochs (n=80). Note C_{xy} is a com-

plex number from which a magnitude and phase relationship can be derived.

3.3 Results

3.3.1 MEA Recordings of OB Slices

Endogenous spiking activity was used to orient the OB slice onto the MEA (Figure 3.1A,B). The large cells in the mitral cell layer produced spontaneously observable spiking waveforms that were easily observed prior to stimulation as well as allowed postexperimental layer identification by plotting the integrated spike range (300-3000Hz) power for all sixty electrodes (Figure 3.1A). The spiking range power overlapped with the external plexiform layer (EPL) adjacent to the mitral cell layer (Figure 3.1A, B top). The electrodes beneath the internal plexiform layer (ITL) and granule cell layer did not record spiking activity (Figure 3.1B bottom).

The integrated gamma band activity (20-70Hz) was generally recorded from electrodes beneath the EPL and mitral cell layer of the OB slice (Figure 3.1C). Endogenous gamma band activity indicated that gamma oscillation power was produced in the EPL/MC layer of the olfactory bulb and spontaneous MC layer activity was sufficient to produce this frequency band of extracellular field activity. The region of the OB slice that exhibited gamma range activity overlapped with the exhibited bulbar spiking activity allowing for the offline OB slice layer identification to follow the plotting RMS voltage of activity from the spiking and gamma band.

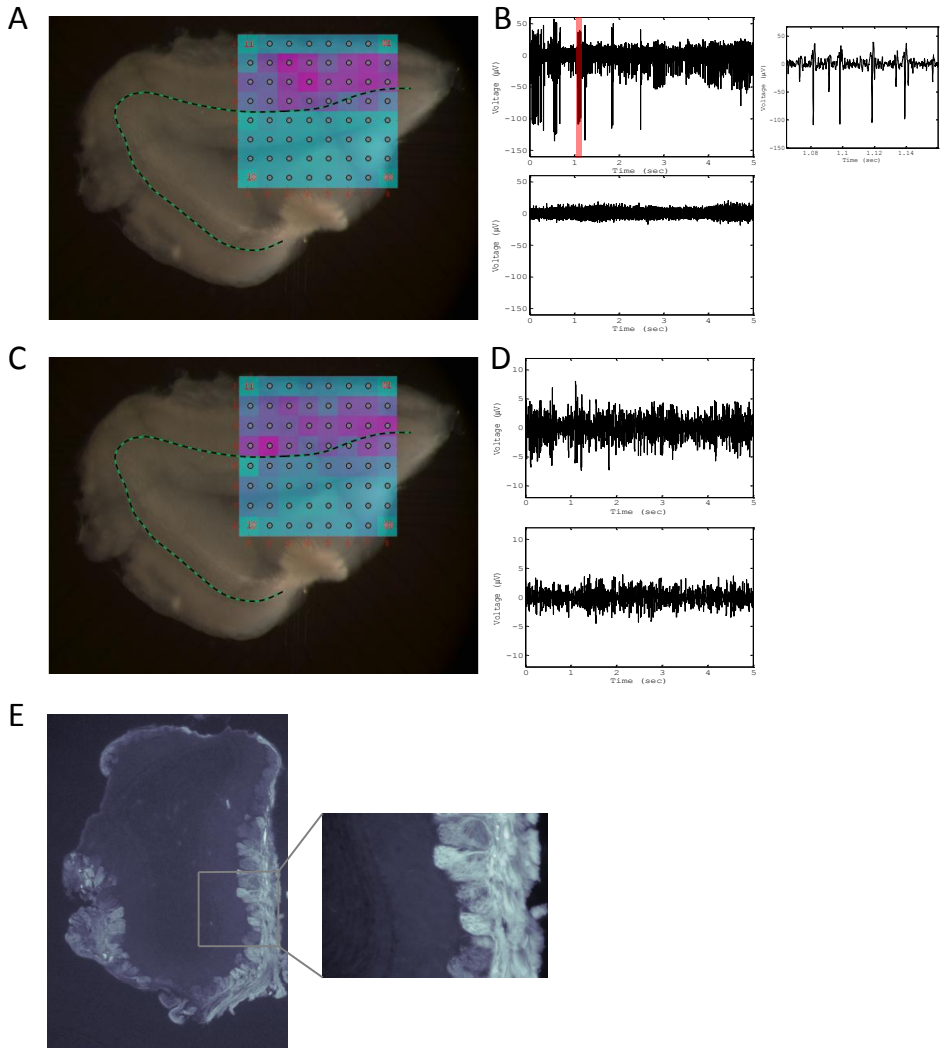

Figure 3.1: MEA recordings from olfactory bulb slices

3.3.2 Pharmacological and Optical Stimulation Induces of Persistent Gamma Oscillations

Cholinergic agonists (Carbachol), metabotropic glutamate receptor group I agonists (DHPG) and group I/II agonists (ACPD), and optical stimulation of ChR2 expressed in the OSN axon terminals induced long lasting oscillations in the

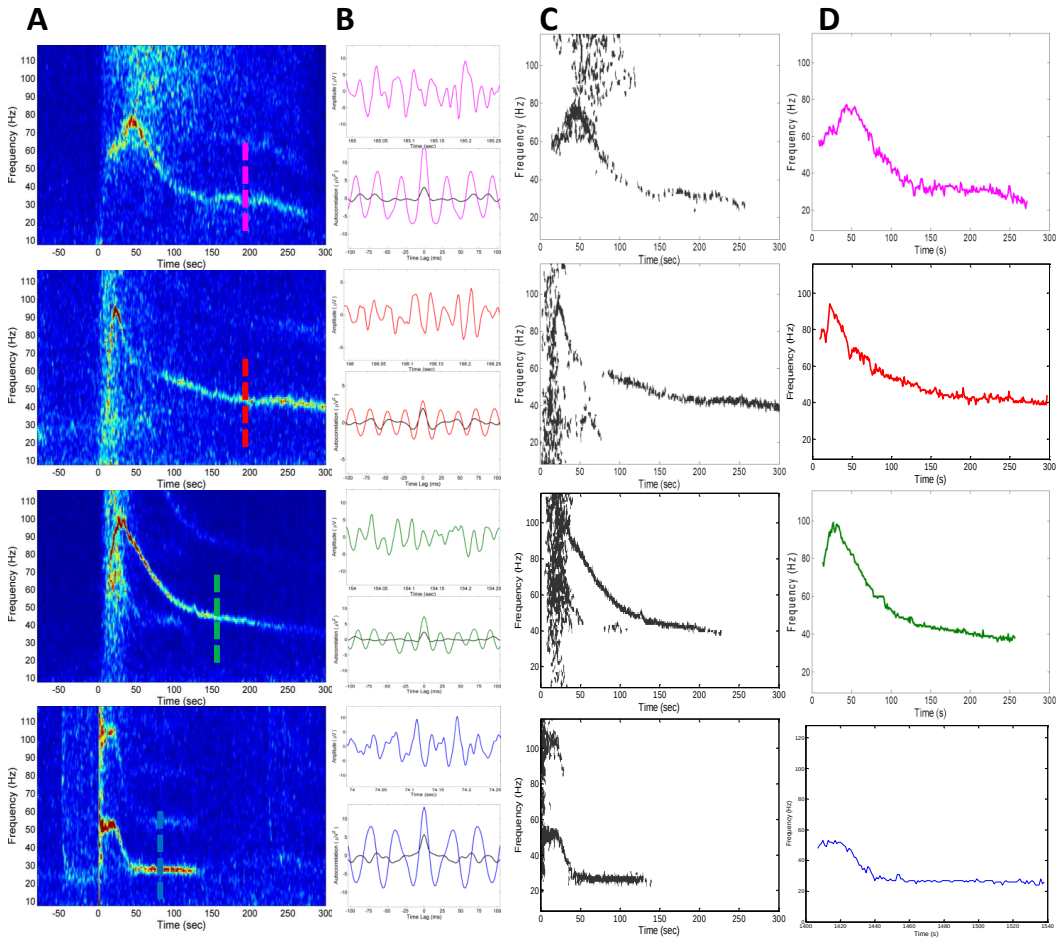

Figure 3.2: Planar electrode recordings and gamma oscillation detection

gamma band (Figure 3.2A, B bottom). Gamma oscillations were first detected by plotting single electrode spectrograms (Figure 3.2A). The gamma oscillations induced persisted up to five minutes. Autocorrelation of shorter time epochs from a single electrode also revealed gamma range side peaks (Figure 3.2B bottom) compared to data recorded prior to stimulation Figure 3.2B top). These data indicate gamma oscillations also occurred on shorter, naturalistic time scales despite the absence of respiration induced theta rhythms.

Gamma oscillations that persisted for minutes were detected from plotted

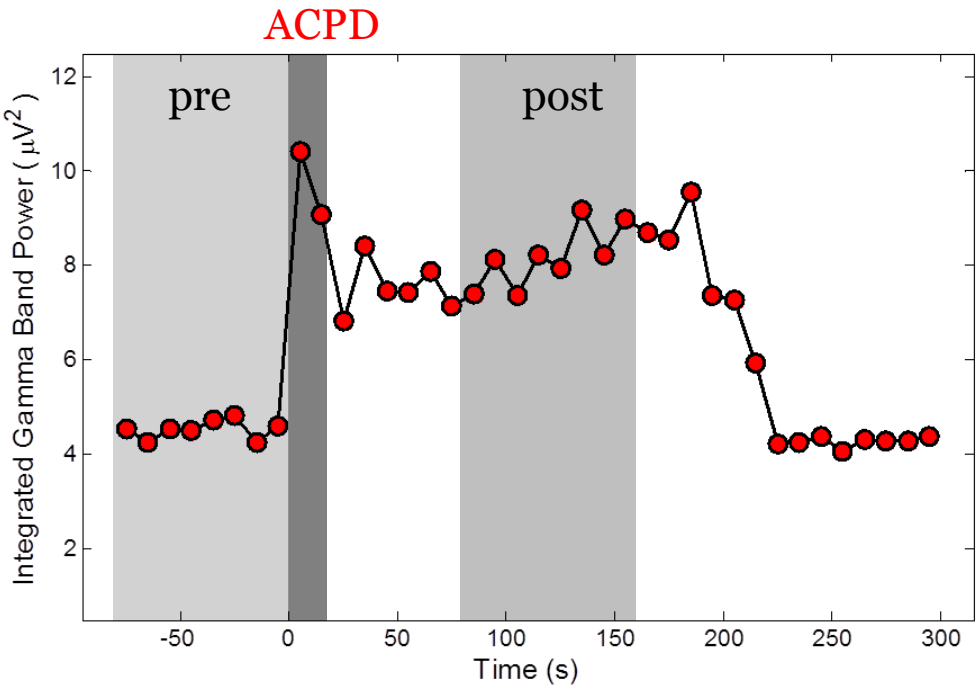

Figure 3.3: Integrated gamma band power across time with ACPD stimulus

spectrograms (Figure 3.2C) as described in Methods. Figure 3.2C illustrates the extracted pixel clusters from the spectrograms in Figure 3.2A. The spectrogram extraction algorithm allowed for partially blind, single electrode gamma oscillation detection. Following the detection of gamma oscillations, spectrograms were plotted and the gamma power trace was traced (Figure 3.2D) allowing the extraction of gamma oscillation statistics (frequency, power, and time).

Traditional definition of gamma oscillations as a simple frequency band (20-70Hz) also showed that the induction of gamma oscillations is an acute affect of the stimulation protocol. Figure 3.3 shows an example of long-term enhancement of the integrated gamma band power under ACPD stimulation and washout of the drug as the power eventually returns to prestimulus level.

3.3.3 Gamma Oscillations are Recorded Across Multiple Electrodes

Gamma oscillations were induced across the OB slice and recorded by multiple electrodes (Figure 3.4A inset). Multielectrode gamma oscillation analysis was performed by overlaying the gamma traces from a single slice (Figure 3.4A). Oscillations chosen for further analysis were considered to have reached a steady state. Steady state oscillations were defined as those oscillations from a single slice that had reached a convergent or ”flatline” frequency (within 5Hz range) and persisted for at least 80 seconds. The highlighted portion of Figure 3.4A illustrates the steady state of the oscillation traces that were compiled in the data pool.

The average frequency of each 80 second oscillation for the separate induction methods was calculated (Figure 3.4B). Carbachol (Figure 3.4B top) produces oscillations that were typically recorded in the 35-40Hz range. The metabotropic glutamate agonist DHPG produced oscillations that were mostly constrained to the 20-25Hz range (Figure 3.4B row 2) while the broader metabotropic glutamate receptor agonist ACPD produced oscillations across a wider range of frequencies (25-30 and 35-40Hz; Figure 3.4B row 3). This trend is more easily visualized in Figure 3.4D. Optical stimulation of glomerular layer expresson of ChR2 produced oscillations mostly in the 25-30Hz range.

The poststimulation integrated power from the stimulated gamma band (20-55Hz) showed that carbachol, DHPG, and ACPD stimulation produced a significant increase in the gamma band compared to endogenous activity prior to stimulation (Figure 3.4C). Conversely, optical stimulation of ChR2 across the

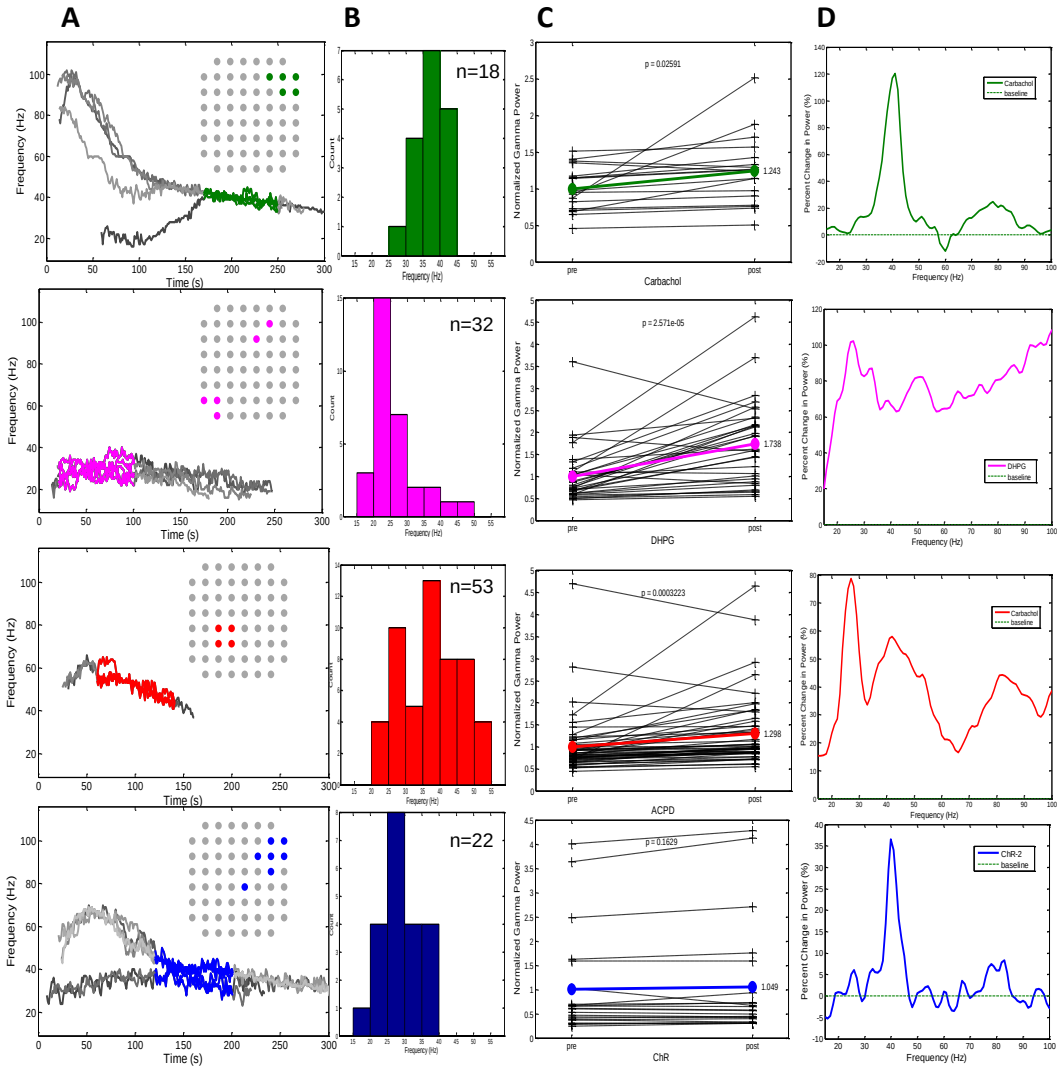

Figure 3.4: Gamma oscillation analysis across electrodes

bulb slice did not produce a significant increase in gamma activity (Figure 3.4C bottom). This result was possibly due to the comparably high amount of endogenous activity present in the analyzed slices. Fast Fourier transform analysis of the ChR2 data pool (Figure 3.4D) that is normalized to the endogenous activity does show a very narrow peak indicating that the endogenous activity recorded from the transgenic OB slices would converge on a very narrow

frequency band following optical stimulation. While the overall power of oscillating network did not significantly change there was a consistent frequency shift.

Similar to the oscillations induced by the optical stimulation of ChR2, application of carbachol also produced a sharp oscillation frequency band following cholinergic stimulation (peak frequency = 40Hz; Figure 3.4D top). On the other hand, mGluR agonist stimulation produced a broad spectral range of steady state oscillations (Figure 3.4D middle rows). DHPG and ACPD produced primary power peaks 27Hz but also produced broad ranges of power across the higher gamma range (40-55Hz). The bulbar stimulation of mGluR receptors also produced long lasting spiking activity that produced an observable artifact in the low passed data in the upper spectral range of the plotted overlaid FFT (70-100Hz; Figure 3.4D middle rows).

3.3.4 Olfactory Bulb Slices have Multiple Regions of Coherent Gamma Oscillations

Broadly stimulating the OB slice with pharmacological agonists and with optical stimulation could induce gamma oscillations across the OB slice as multiple gamma band oscillations were recorded simultaneously with multiple electrodes. Figure 4.5A illustrates two independently induced gamma oscillations 600μm apart. Gamma oscillation coherence in the intact bulb is an active area of research (Kay et al., 2009 [**?**]). The simultaneous oscillations illustrated in Figure 4.5A had separate frequency components at which neighboring oscillations would cohere producing independent coherence magnitude signatures

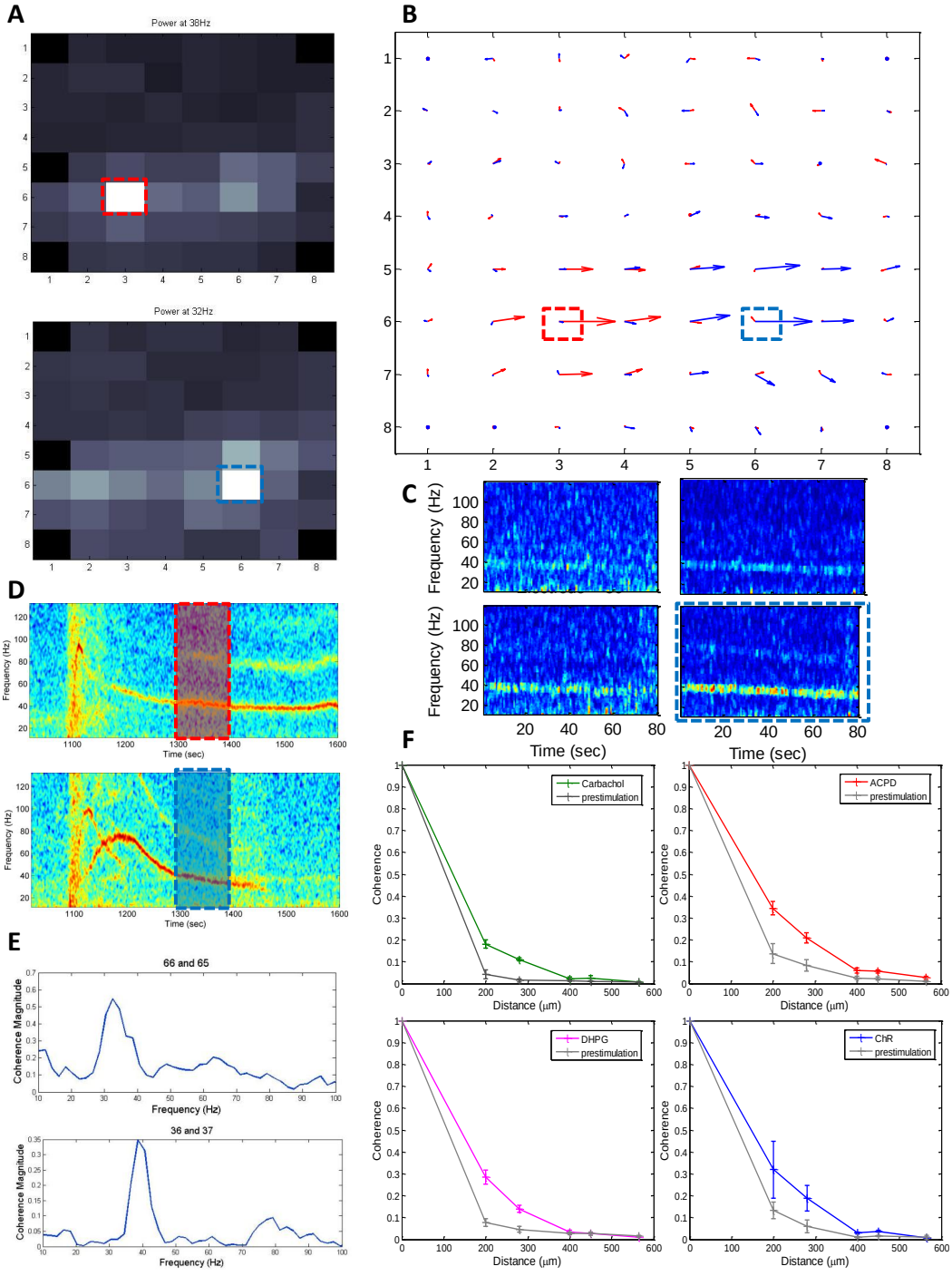

Figure 3.5: Coherence analysis of gamma oscillations across OB slice

(Figure 4.5B). The coherence magnitude plots for each reference electrode in Figure 4.5A (top = e36; bottom = e66) revealed the frequency with peak coherence magnitude to be used for further coherence analysis. Plotting the power at the reference frequency from each electrode illustrates the existence of independent regions centered around the reference electrodes in Figure 4.5A (Figure 4.5C). We used this data to produce quiver plots (Figure 4.5D) to determine the span of the coherence pockets centered around our reference oscillations. These plots show that gamma oscillations in the OB slice can span across multiple electrodes (200μm between each electrode) with multiple, independent regions of coherence simultaneously occurring. To further quantify the spread of coherence across the bulb slice, the average coherence magnitude from electrodes adjacent to the oscillation epicenter for each oscillation induction category were calculated from quiver plot data. This data shows that gamma oscillations were coherent up to 280μm across the MC/EPL in OB slices within all of our induction groups.

3.4 Discussion

Overall, the MEA provided an advantageous method to investigate the complex, dynamical synchronization properties of olfactory bulb circuitry. How information is synchronized across space as well as transformed through multiple layers of neural circuitry is an active area of research. The large recording surface area of the MEA increased the ability to detect the electrical signatures of synchronized neural activity and provided a means to analyze multiple, simultaneous waveforms allowing further insight into neural assembly information propagation.

The resulting OB slice data support the conclusion that a variety of stimulation mechanisms can induce gamma oscillations in the OB slice consistent with the gamma band range found in other studies using OB slices (Friedman and Strowbridge, 2003 [28]; Galan et al., 2006 [30]; Schoppa, 2006 [74]). One particularly interesting result is the presence of endogenous gamma oscillations prior to stimulation readily recorded across the bulb slices. These typically lower frequency oscillations existed endogenously and were recorded prior to stimulation indicating that spontaneous activity from tufted cells located in the glomerular layer or mitral/tufted cells in the mitral cell layer may provide sufficient activity to drive gamma oscillations in the absence of odor stimulation.

The induction categories, cholinergic agonists, metabotropic glutamate receptor agonists, and tetanic stimulation, were ported from the stimulation methods used to induce synchronized activity in hippocampal gamma oscillation studies (Whittington, et al., 2000 [88]). Our results indicate there may be similar diversification of these receptor types across the circuitry that coordinate biophysical gating of information in the hippocampus. The oscillations induced by these stimulation methods in the OB slices persisted up to 5 minutes. These long lasting oscillations occurred widely across the OB slices and presented complex coherence relationships. The presence of fragmented pockets of coherence that can extend up to 280μm across the bulb slice indicate the existence of circuitry capable of cohering oscillating regions across the intact OB. Sister mitral cells can extend up to 70-100μm across the bulb. These results support the argument that our bulb slice oscillations extended across an estimated 3-4 glomerular column widths. The identification of the neural circuitry that functions to coordinate synchronized activity within the olfactory bulb remains to be identified.

The presence of parallel coherence pockets has been predicted to exist in biological models supported by data from computational models of reduced connectivity across independently oscillating neural assemblies (Bazhenov et al., 2008 [62]). We here hypothesize the fragmented regions of coherence recorded across the MEA are an artifact from slicing the OB and are a result of reduced lateral connections providing the necessary mechanisms to coordinate multiple oscillating regions. Limiting the coupling between oscillating areas in a variably stimulated subsection of neural circuitry could plausibly produce an increased drift in oscillation frequency resulting in variable regions of coherence profiles.

CHAPTER 4

SCALABLE NEURAL RECORDING: CMOS MICROELECTRODE ARRAYS FOR SLICE RECORDING

4.1 Introduction to CMOS microelectrode arrays

How networks of neurons work together to perform complex tasks and propagate information is an active area of research. To effectively study neural ensembles, neurophysiologists employ microelectrode arrays (MEA) to simultaneously acquire electric field activity across relatively large areas of tissue. The advantages of increasing the number of recording electrodes include: 1) the possibility for a greater number of single cell recordings and 2) spatially broad analysis of local field potentials (LFPs) that can provide insights into how and why neuronal ensembles synchronize their activity. MEA data are an invaluable resources for computational neuroscience and for the study of biological learning, acute and chronic drug effects, and physiological disorders.

Conventional planar MEAs are noninvasive and function by recording electrical extracellular activity with metal electrodes (Figure 4.1). Signals detected at the electrode interface are amplified, passed through a low-pass filter, and then digitized. MEAs are well suited for extended *in vitro* studies using either tissue preparations, such as retinas or brain slices [47], or neuronal cell cultures [27]. Most commercially available MEAs are fabricated on passive, patterned substrates which use off-substrate amplifiers. While these MEAs typically have between 60-256 electrodes, passive routing limits the electrode scalability. CMOS MEAs can achieve a higher spatial resolution and a large effective recording area by multiplexing channels onto fewer wires. Furthermore, they can include

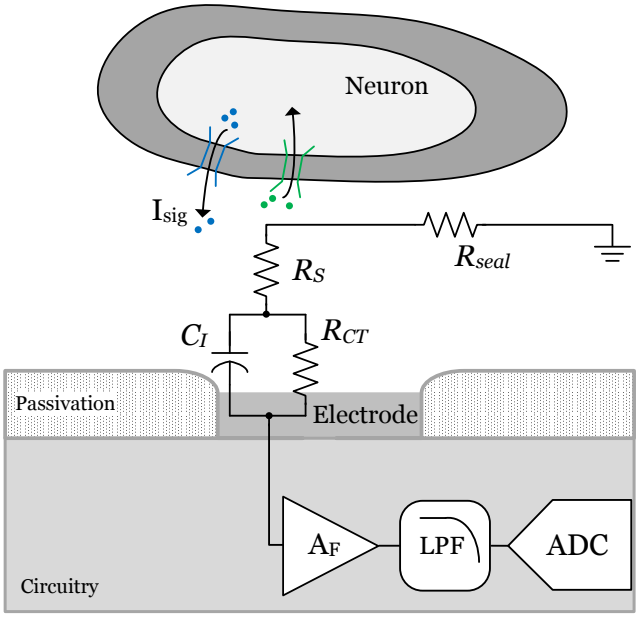

Figure 4.1: Concept of planar CMOS MEA recording

additional modalities, such as optical recording [47], electrical stimulation [27], or temperature sensors [40].

Another effective way to study neural networks with MEAs is in conjunction with optical stimulation. In visual neuroscience, computer-generated patterns of light can be used to stimulate retinal photoreceptors while electrically recording activity from the ganglion cells, the retinal output. Likewise, optogenetics allows researchers to create light-sensitive cells in other biological systems to understand the way they process information. For example, neuronal cells expressing the light-reactive protein channelrhodopsin-2 can be excited optically, circumventing the need for either electrical stimulation which creates large recording artifacts, or chemical stimulation which is nonspecific [10]. Optical stimulation can be spatially and temporally precise, cell-specific, and contact-free. Therefore, optical stimulation allows researchers to control the inputs of a

neural system while an MEA allows them to record multiple intermediate and output nodes of the network.

Passive MEAs are often built on transparent substrates allowing for visual correlation of electrode location, tissue, and stimulus. Active MEAs on silicon substrates are opaque and severely limit this capability. Furthermore, CMOS MEAs can still suffer from some of the same scalability limitations of passive MEAs. State-of-the-art CMOS arrays typically tradeoff between electrode density, noise performance, and sampling rate. Therefore, the aim of this work is provide researchers with an MEA that overcomes the limitations of both traditional passive and active MEAs for slice research. The presented MEA has a high spatial resolution, a large recording area, and integrated angle-sensitive pixels (ASP [85]) to correlate optical stimuli with recorded activity. Additionally, the sensor array requires little post-CMOS processing for biocompatibility and can record from all sensors simultaneously.

4.1.1 Methods

The design was dictated by the characteristics of extracellular biopotentials recorded from olfactory bulb slices. A commercially available MEA (MEA60-Up-System, MCS GmbH) with 60 TiN electrodes was used to record extracellular fields. The electrodes had a diameter of 30μm and a pitch of 200μm. Slices were horizontally taken with a vibrating microtome with a nominal thickness of 300μm. During recording, slices were perfused with artificial cerebrospinal fluid (aCSF) containing (in mM) 125 NaCl, 25 $NaHCO_3$, 1.25 NaH_2PO_4, 25 glucose, 3 KCl, 1 $MgCl_2$. The solution was heated to 34°C using an in-line heater

and oxygenated with carbogen (95% O_2, 5% CO_2). Slices were held down by a chrome harp with nylon netting.

4.1.2 Extracellular Biopotentials

According to the literature, discernible extracellular biopotentials typically range in amplitude from 10μV to about 5mV, with power in the 10-200Hz band for LFPs and 200Hz-5kHz band for action potentials [38]. Figure 4.2 shows the measured power spectrum of endogenous slice activity with several electrodes (n = 13). The measured spectrum was calculated by 1) averaging power spectra from a 50s long single electrode recording with 1s epochs, then 2) averaging that power spectrum with n other electrodes. The power spectrum reflects contributions from both LFP and action potentials. The LFP power peaks around 45Hz, which corresponds to the gamma band in slice (30-70Hz [34]). Gamma oscillations are ubiquitous in the brain and are indicators of synchronous cellular activity. While gamma oscillations can occur endogenously, they can also be induced chemically with glutamate receptor agonists. A spectrogram of an induced gamma oscillation is shown in Figure 4.3. Oscillations were also induced by optical stimuluation of slices from transgenic mice expressing channelrhodopsin-2 in the olfactory sensory neuron axons. To utilize the temporal and spatial precision of optical stimuli, the proposed MEA uses photopixels to correlate light stimuli with recorded electrical activity.

Neural acquisition systems for *in vivo* recording deal with large voltage offsets from LFP, and often separate the LFP and spike bands to alleviate the dynamic range requirement of the ADC [67]. However, in this experiment the

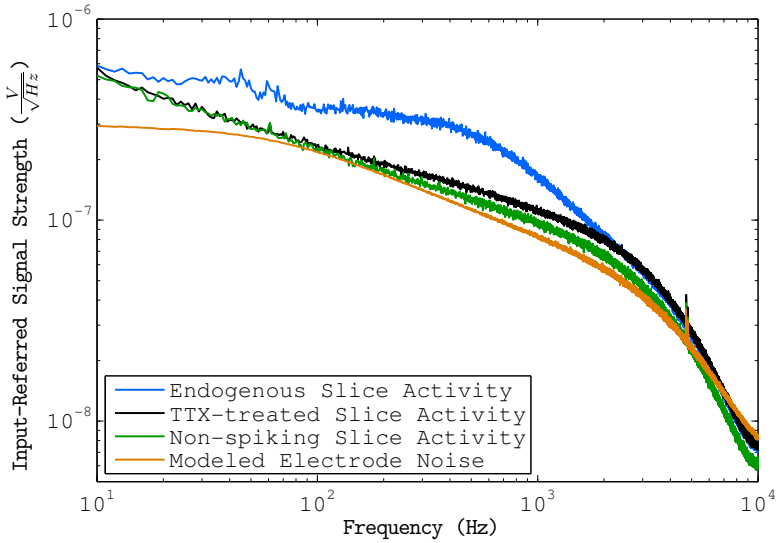

Figure 4.2: Power spectra from MEA slice recordings

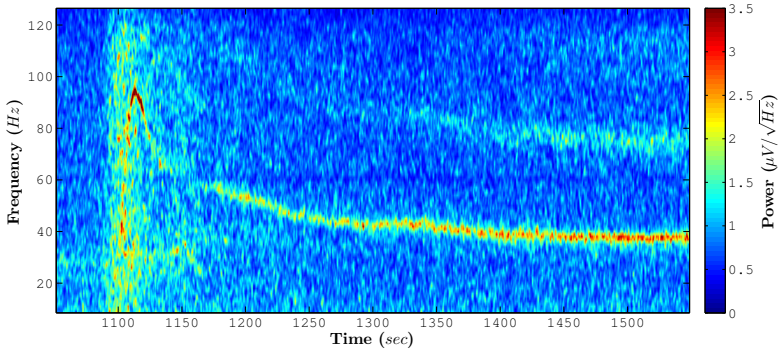

Figure 4.3: A spectrogram of an induced gamma oscillation

largest recorded biopotentials were on the order of 200μV_{pp}. Furthermore, the frequency content of LFP and spiking can overlap. The resultant spectral contamination means that traditional low-pass filtering methods to remove spike contributions from the LFP band may give rise to unwanted spurious correlations, problematic when studying the causal relationship between LFP and spiking. Spike removal algorithms have been shown to effectively remove spike contributions from the LFP band, assuming the recorded signal is wideband

[92], [17]. Thus, the proposed MEA does not separate the LFP and spike bands on chip since more sophisticated filtering is required.

4.1.3 Background Noise Level

Recording neural signals requires low-noise instrumentation. To understand the background noise level associated with recording from slice, a sodium channel blocker, tetrodotoxin (TTX), was applied during the recordings (n = 13). This yields an estimate of the recording noise level because the spiking and oscillations have been blocked. An additional background level was obtained by averaging spectra recorded from electrodes covered by slice tissue without discernible spikes or oscillations (n = 9). The total integrated noises were 7.2μV_{rms} and 6.5μV_{rms}, respectively. With microelectrodes, the dominant noise source in extracellular recording typically arises from the electrode-electrolyte interface, not the recording circuitry which is often well controlled [90]. Using methods presented in [49] and [24] for the circuit model in Figure 5.1, we estimated the charge transfer resistance (R_{CT}) of the interface to be 4.7MΩ, the parallel interface capacitance (C_I) to be 495pF, and the spreading resistance (R_S) to be 12kΩ. Neglecting the Warburg impedance, the estimated noise from the interface was 4.9μV_{rms}. The recording circuitry had little effect on the overall noise level (2.6μV_{rms}), which was measured by shorting the input of the recording amplifier to ground (n = 3). In terms of frontend amplifier design, circuit noise levels much below 2.4μV_{rms} contribute little to the overall noise level and are likely overdesigned, consuming unnecessary power or area. Circuit noise above 5.3μV_{rms}, however, is likely to be the dominant noise source for electrodes with similar geometry.

4.1.4 Spatial Spread

Measured persistent epochs (> 5 min) of gamma oscillations (30-70Hz) spanned over 600μm laterally while spiking activity from individual cells typically spread less than 100μm. This is comparable to measurements from hippocampal slices [26] and modeled results for synchronous network activity [57]. The spatial reach of gamma oscillations was determined by finding the coherence between the electrode with the highest gamma power and all other electrodes, where the coherence magnitude between two electrode recordings in the frequency domain is given by:

$$C_{xy}(\omega) = \frac{\left|\langle P_{xy}(\omega)\rangle_n\right|^2}{\langle P_{xx}(\omega)\rangle_n \cdot \langle P_{yy}(\omega)\rangle_n}. \tag{4.1}$$

P_{xx} and P_{yy} are autocorrelations and P_{xy} represents the cross-correlation between the two electrodes. Note that coherence for every 1Hz frequency bin was calculated over n epochs of 1s, where n = 150. Figure 4.5 shows an example spatial coherence profile. In theory, adjacent electrodes exhibit no coherence in the absence of synchronous activity ($\mu_0 = 0$) while oscillatory activity spanning multiple electrodes will be highly correlated. The variance of coherence was calculated using 58 electrodes and 20 frequency bins (21-40Hz) for a total of 1160 measures of coherence. Based on the spatial expanse of LFP, a useful planar MEA should have an active area larger than 1mm on a side to accommodate large synchronous oscillations from an entire slice and have an electrode pitch of less than 100μm to fully sample extracellular action potentials.

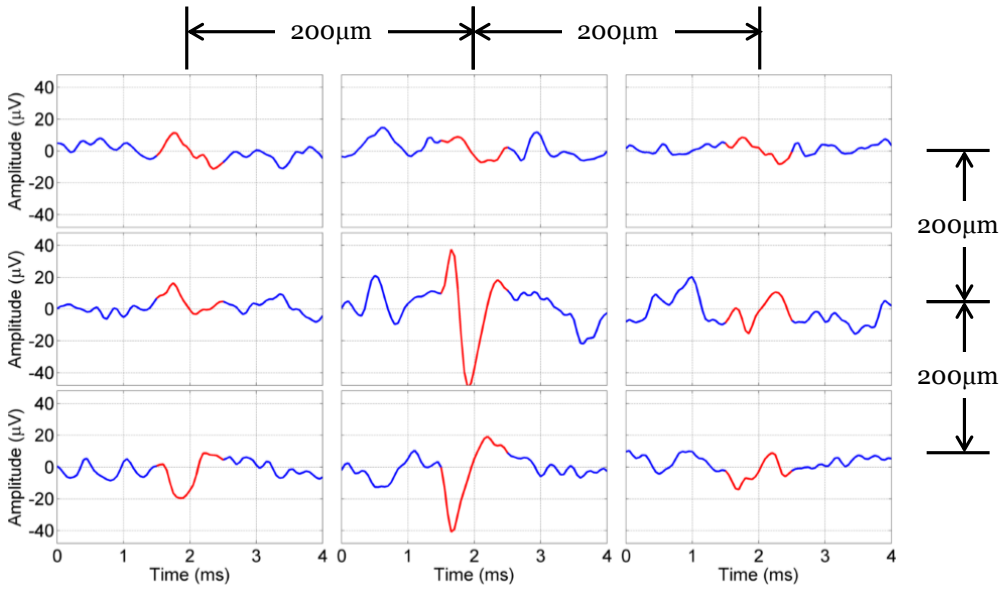

Figure 4.4: Spatial spread of action potentials

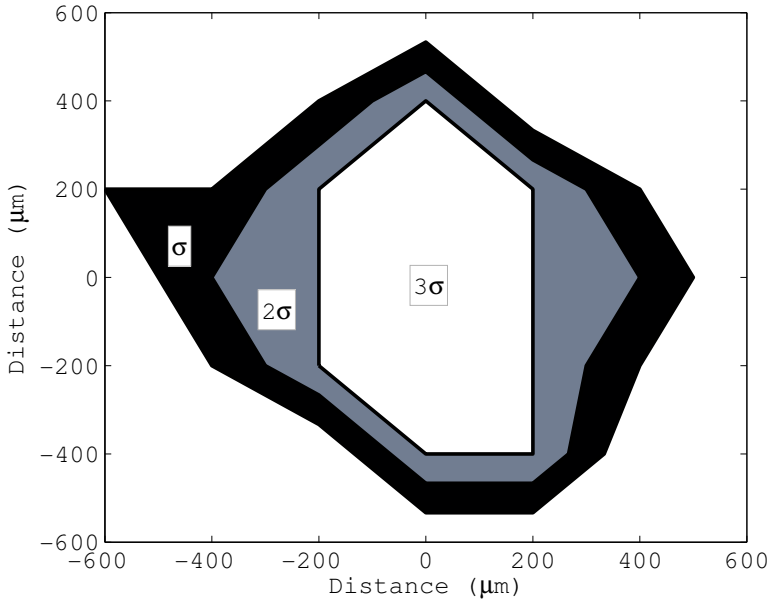

Figure 4.5: Spatial coherence profile of a 33Hz gamma oscillation

Table 4.1: Design Requirements of the CMOS Sensor Array

Specification	Requirement
Signal Bandwidth	10Hz - 3kHz
Sampling Rate	10kHz
Circuit Noise	2.4μV_{rms} - 5.3μV_{rms}
Dynamic Range	>30dB
Electrode Pitch	<100μm
Spatial Extent	>600μm

4.1.5 Design Requirements

Table 4.1 provides a summary of the design requirements for the CMOS sensor array derived from experimental MEA recordings. While previous work has achieved sub-cellular electrode pitch (<10μm) and low-noise amplification (2.4μV_{rms}), the array record from a static selection of 126 electrodes simultaneously [27]. In principle, these sites can be multiplexed faster than the bandwidth of the neural signals, allowing multiple sites to be measured each sample cycle. The difficulty with high-speed multiplexing before amplification is that noise from the interface is not filtered and is therefore aliased into relevant signal bands. Other work has used electrode-level amplification to increase the number of recording channels. [8] provides pixel-level amplification with 4,096 small pitch electrodes (42μm) but sacrifices noise performance (11μV_{rms}). Another high-density array [53] has an even finer sensor pitch (7.8μm) and more sensors (16,384), but with a much higher noise level (70μV_{rms}). Therefore the primary design challenge of CMOS MEAs is designing a low-noise amplifier with a very small area and with a scalable data read-out.

4.2 A 768-Channel CMOS Microelectrode Array with Angle Sensitive Pixels for Neuronal Recording

This section discusses the first-generation design of a CMOS microelectrode array in 130nm CMOS (published in [46]). The array consists of 768 low-noise recording channels integrated with angle-sensitive pixels (ASPs, [84]). The first-generation array is highly scalable due to electrode-level digitization with serial data stream-out. The front-end amplifiers use chopping to reduce flicker noise and achieve an input-referred noise of 4.1μV_{rms} over a 3.6kHz bandwidth while occupying an area of only 800μm^2. Digitization is performed by using a distributed 10-bit ramp ADC that samples every sensor site at 10kHz. The electrodes have a 50μm pitch and were plated with platinum to increase the interface capacitance and ensure biocompatibility. Sensor array functionality was demonstrated by refocusing a lenless image and recording neural spiking and local field potentials from a mouse olfactory bulb slice.

4.2.1 System Description

The hybrid sensor array consists of 768 metal electrodes with local amplification and digitization and 2,048 photopixels. The MEA was fabricated in a standard 130nm CMOS process with an active area of 1.6mm by 1.6mm (Figure 4.6). An input clock (F_{in} = 40MHz) feeds a 12-bit counter in the digital core which then generates the control signals and clocks for the rest of the chip. Each sensor site is sampled at a rate of 10kHz as determined by the MSB of the counter, which is the main clock rate divided by 2^{12}. Figure 4.7 shows the timing diagram of the

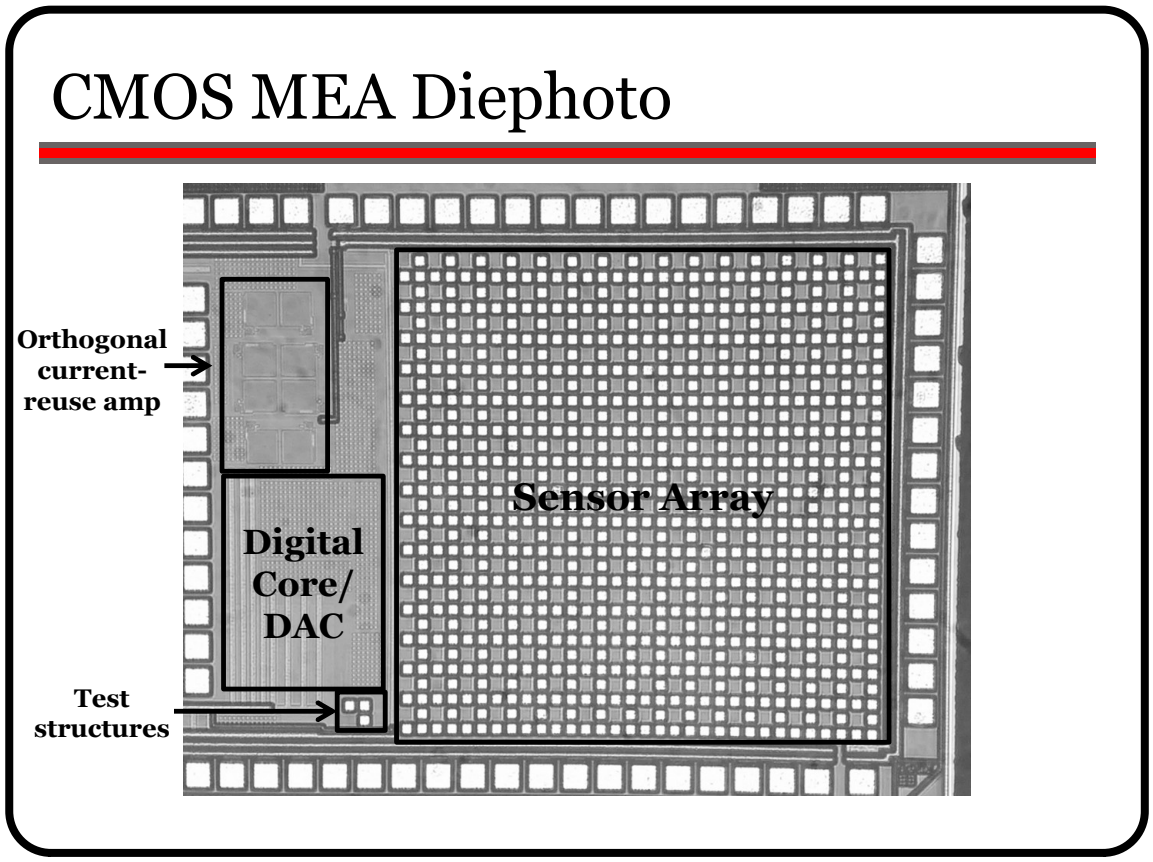

Figure 4.6: Die photograph of the sensor array

array. Sensor sampling is interleaved so that data can be output continuously at a rate equal to the main clock. There are four data output channels, each of which is fed by its own 640-bit shift register. Half of the sensors are sampled simultaneously while the other half of the sensors load and then shift data off the chip. A global 20kHz ramp signal is used as an ADC reference voltage at every sensor. The ramp is generated by a PFET current DAC controlled by the counter with a resistive load to translate the output current into a voltage. Since control signals are generated globally and the outputs are digitized, the design is highly scalable. The limiting factor of this design was the available die area.

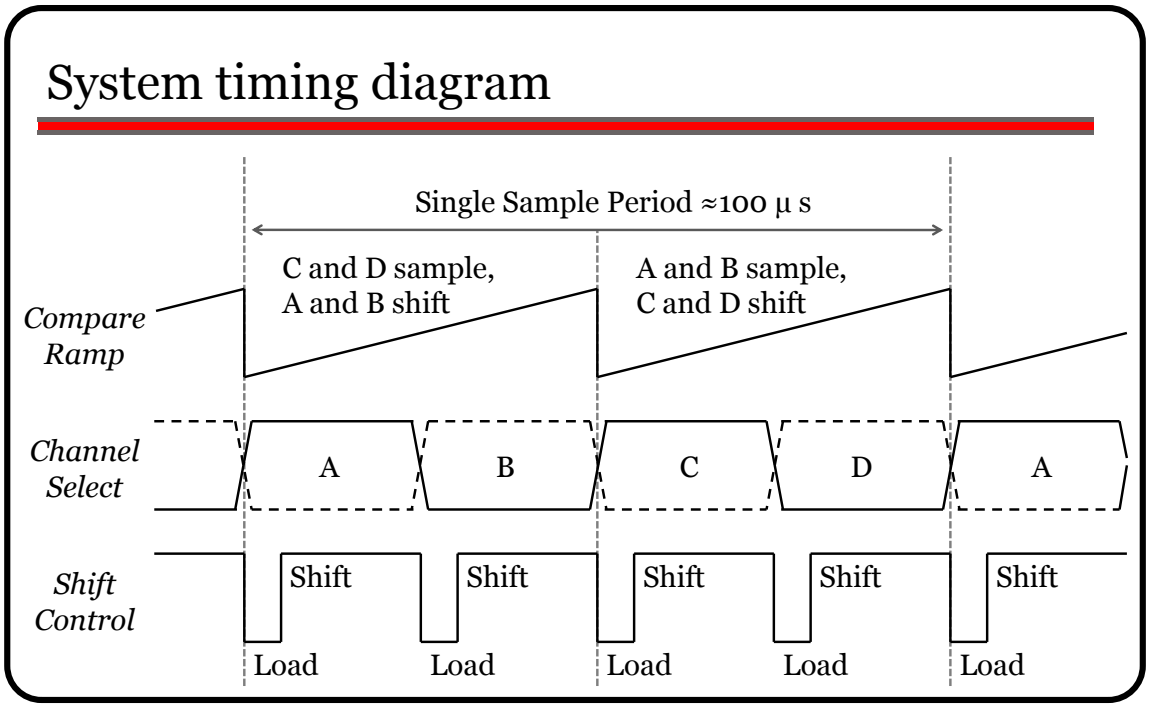

Figure 4.7: Timing diagram of the array control signals

4.2.2 Four-Sensor Subunit

Figure 4.8 shows the four-sensor subunit. Each subunit is comprised of three recording electrodes (**A**, **B**, **C**) and a set of 8 ASPs (**D**). Each site is allocated an area of 50μm by 45μm with an additional 50μm by 5μm of each sensor site used as part of the global shift register, which streams out the locally stored data.

The backend for the recording electrodes and the ASPs are identical. The comparator digitizes the analog signal from the sensor by comparing it to the global ramp signal. Each step of the ramp corresponds to a 10-bit value generated by the counter in the digital core. When the ramp signal is larger than the amplified signal, the comparator clocks the latch, storing the 10-bit number lo-

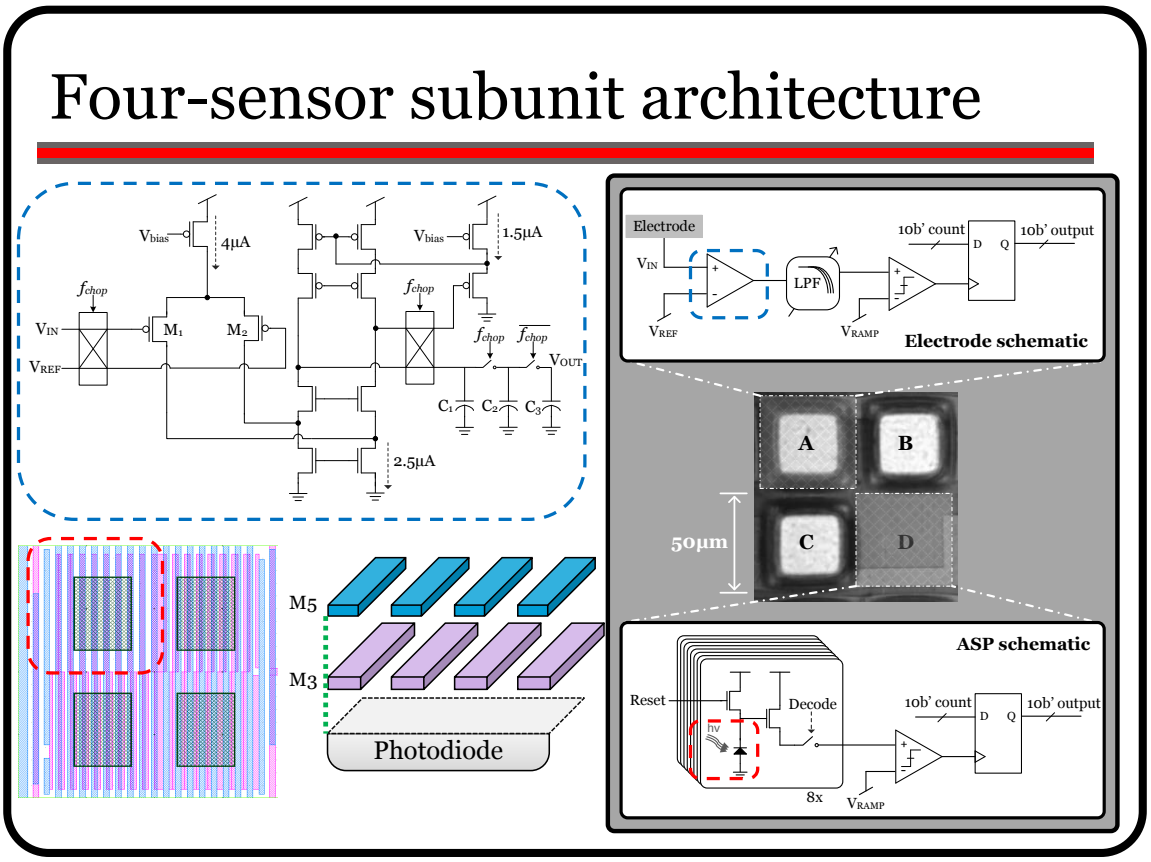

Figure 4.8: Architecture of four-sensor subunit

cally, and then shuts down to conserve power and reduce switching noise until the next sampling cycle.

4.2.3 Frontend Amplifier

Each recording electrode site contains an amplifier, a switched-capacitor low-pass filter, a comparator, and a 10-bit latch as shown in Figure 4.8. The metal electrode interface is defined by a passivation opening over the top metal and is formed during the standard CMOS fabrication process. V_{REF} is the applied

reference potential of the electrolyte bath and is common to all the amplifiers of the array. Figure 4.8 shows a schematic of the folded cascode front-end amplifier. Traditional neural amplifiers use large input transistors to reduce flicker noise and large capacitors (20pF) to block electrochemical offsets [38]. To reduce area in this design, the front-end amplifier uses chopping modulation to reduce the flicker noise of the amplifier and bias the input transistors. In simulation, the total input-referred noise voltage of the amplifier over a 10kHz bandwidth is reduced by a factor of 6.6 when chopping modulation is used. The size of the input transistors are small to maintain a high input impedance (>20MΩ at 1.25MHz). To achieve the same performance without chopping modulation, the input transistor area would need to be increased by a factor of 43.

Low-pass filtering is needed before sampling in order to prevent aliasing of high-frequency noise from the tissue, electrode-electrolyte interface, and the amplifier. Rather than using a large load capacitor to pull the amplifier output pole to below half the sampling frequency, a much smaller MOS capacitor (C_1) was used to prevent aliasing of high-frequency chopping artifacts, followed by a switched-capacitor low-pass filter ($\omega_L = f_{chop}C_2/C_3$). This significantly decreases the required area to get the desired pole without lowering the amplifier current, which would increase thermal noise. Furthermore, the chopping frequency controls the low-pass frequency corner. Note that the chopping frequency is controlled directly by the main clock so that the comparator always samples with the same phase. Another advantage of switched-capacitor filters is that they are relatively process invariant since their corner is set by the switching frequency and a ratio of two nearby capacitors. The amplifier and low-pass filter occupy an area of $800\mu m^2$, about 25 times smaller than neural amplifiers in MEAs with similar performance [7].

4.2.4 Angle-Sensitive Pixel Design

In place of a front-end amplifier, one-quarter of the sensor sites (**D** from Figure 4.8) contain eight distinct ASPs. ASPs, first introduced in [84], are devices which are sensitive to the angle of incident light and have been used to localize multiple fluorescent sources in 3D space [85] and perform post-capture computational refocus of visual scenes [86]. ASPs were implemented rather than standard photopixels because they can provide a more complete description of the light field. This is useful for understanding the scattering of stimulus light in tissue or localizing external electrodes with respect to the array. In general, ASPs use two CMOS metal gratings over a photodiode where the pitch and height between the gratings define the angular selectivity. The lower grating is used to block or pass the periodic intensity pattern generated by light striking the top grating as a function of its lateral offset. In this design, the ASPs employ two local, stacked diffraction gratings on CMOS metal layers 5 and 3 directly over a p-implant/n-well photodiode. At each sensor site there are two orientations of top diffraction grating (vertical and horizontal), and four types of bottom grating offset relative to the top grating ($\alpha = 0$, $\frac{\pi}{2}$, π, or $\frac{3\pi}{2}$), for a total of eight ASP variants. α defines the angle of peak photodiode response. Since the sampling rate needed for an ASP is much lower than an electrode, only one of the eight ASP subtypes are digitized each time for an effective sampling rate of 1.25kHz.

4.2.5 Packaging

A major difficulty with using CMOS sensors is ensuring biocompatibility and protecting wirebonds from the electrolyte solution. Most post-CMOS packaging techniques use lithographically defined epoxy or patterned PDMS to encapsulate wirebonds [16]. To simplify post-processing, epoxy was applied under a light microscope without a mask and used to encapsulate the wirebonds and define a well around the active area. The top metal of this process is aluminum, which corrodes easily in saline [60] and is also cytotoxic [20]. To prevent the electrodes from corroding, the electrodes were electroplated with platinum. Platinum is nontoxic and also decreases electrode impedance. Electroplating was performed by filling the well with a platinizing solution (chloroplatinic acid, lead acetate, hydrochloric acid). 1.5V was applied to a platinum counter-electrode to pull current through the electrodes.

4.2.6 Measurement Results

Figure 4.9**A** shows the measured transfer function of the front-end amplifier with two different chopping frequencies. The amplifier consumed a total of 6μA from a 1.5V supply. The amplifier's bandwidth and midband gain changed from 3.6kHz and 43.5dB with a 1.25MHz chopping frequency to 2.3kHz and 43.9dB with a 700kHz chopping frequency. The total measured input-referred noise was 4.1μV_{rms} over the 3.6kHz bandwidth, corresponding to a NEF of 6.5 [78]. The ramp ADC achieved an ENOB of 8.7 bits with a 512mV_{pp} ramp, which corresponds to a resolution of 7.9μV at the input (f_s = 10kHz, f_{in} = 100Hz). The total power consumption of the sensor core was about 13.3mW, with 55% of the

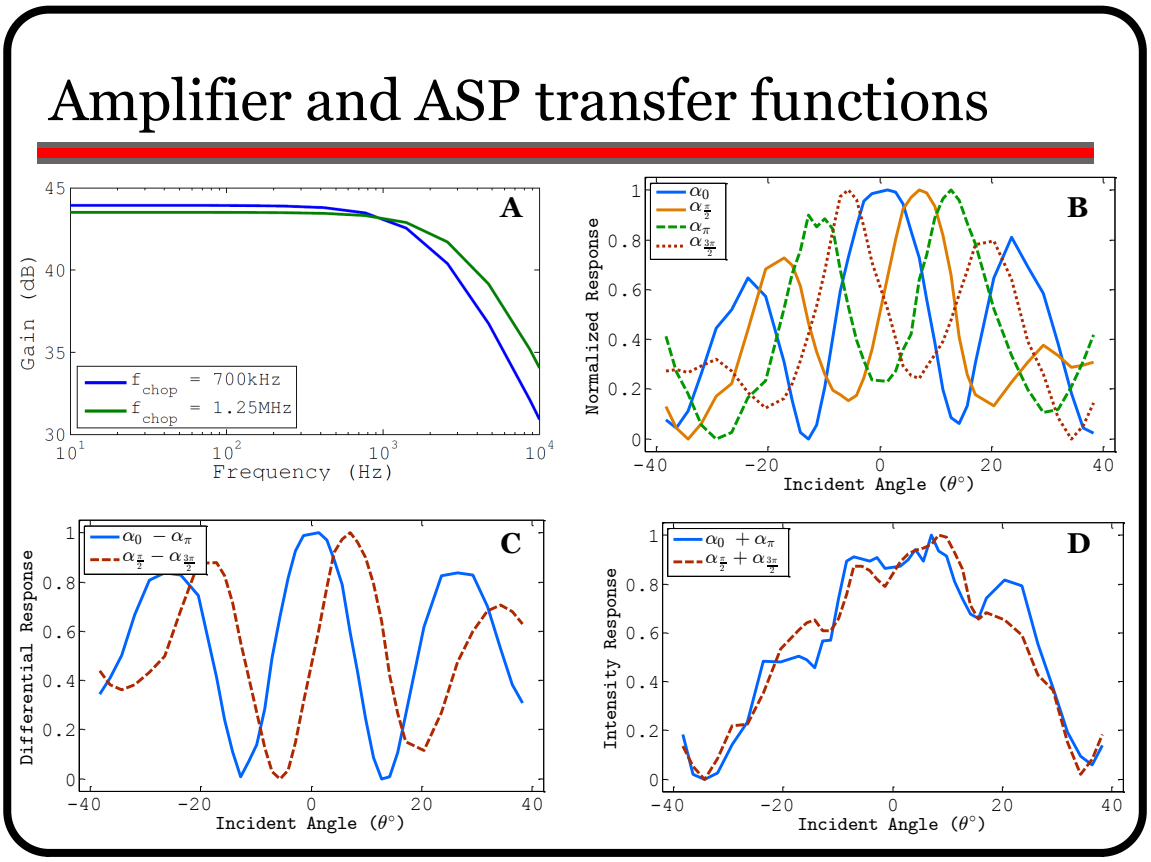

Figure 4.9: Measured frontend amplifier and ASP transfer functions

power consumed by the frontend amplifiers.

Figure 4.9**B** shows the normalized response of four ASP subtypes to changes in incident angle. The light was generated by a 470nm LED which was kept at a fixed distance and rotated by a micromanipulator around the sensor array. Note that the output of an ASP is a function of both intensity and incident angle, which can result in ambiguity between a bright source at a blocked angle or a dim source at a passed angle. This ambiguity is resolved by taking the difference between complementary ASPs, as shown in Figure 4.9**C**. The sum of the complementary ASPs encodes the intensity of the incident light (Figure 4.9**D**).

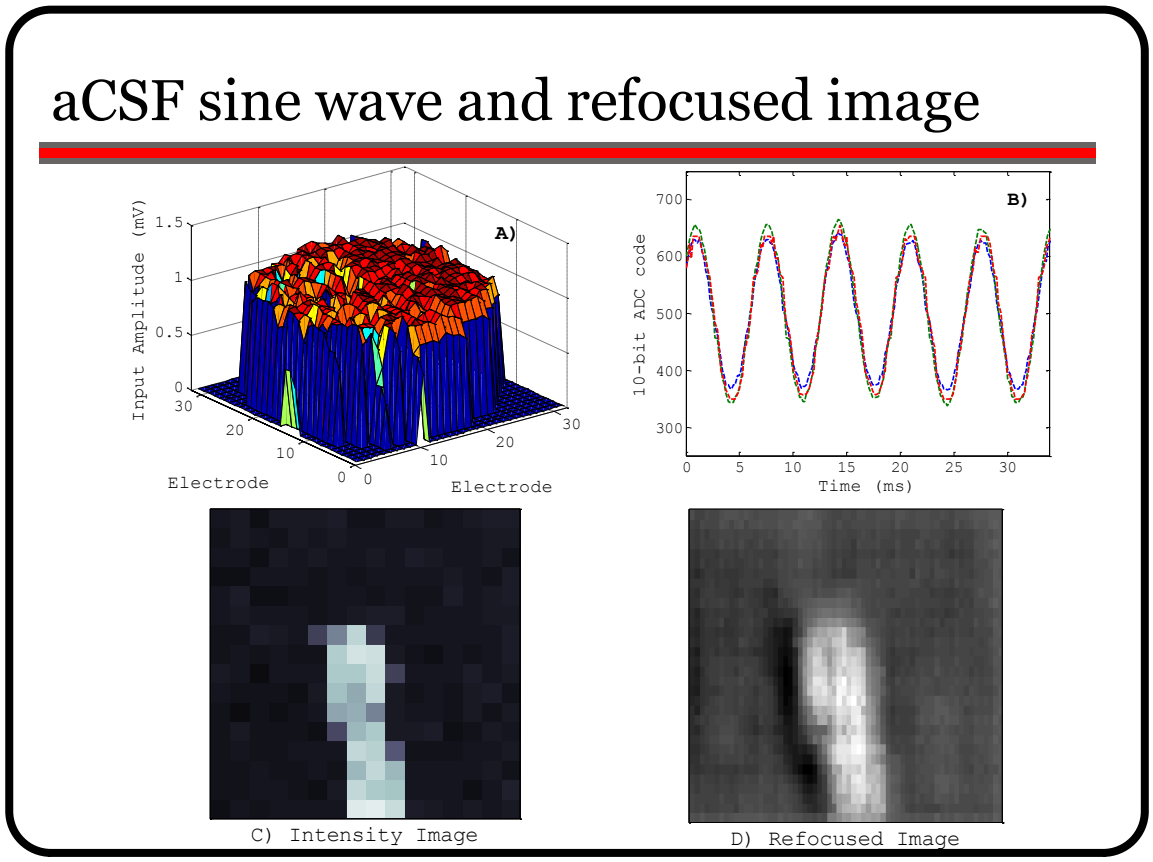

Figure 4.10: sinewave test and refocused image

Since each front-end amplifier operates open loop, there is a slight gain mismatch between electrodes. To calibrate and demonstrate array functionality, the well was filled with aCSF and stimulated with a 1mV sine wave through a silver chloride wire. Figure 4.10A shows a three-dimensional map of the input-referred rms voltage recorded by the array. The average gain across the array was 43dB with a standard deviation of 1.8dB. Figure 4.10B shows a 35ms overlay of three adjacent electrodes from the experiment in Figure 4.10A.

To demonstrate the functionality of the optical read-out, a 250μm diameter platinum reference electrode was fixed 200μm above the sensor array. First, an

intensity image (sum of complementary ASP subtypes) was captured using an LED illumination source. Then, using synthetic refocus techniques from [86], the image was computationally refocused in software as shown in Figure 4.10. Refocusing uses the information of the difference between complementary ASP subtypes and then convolves them with scaled, oriented Gabor filters.

To confirm that the sensors could provide sufficiently robust, low noise recording across the full range of biologically relevant signal bands, they were used to record both LFP and spiking activity. The tissue preparation was identical to that used in Section II. A power spectrum of oscillatory activity from a 300μm thick mouse olfactory bulb slice is shown in Figure 4.11. This data, unfiltered in software, shows an oscillatory peak at the gamma band and a sharp peak at 60Hz due to line noise. The line noise enters the system through the gravity-fed perfusion system and is not inherent to the circuitry. Gamma oscillations arise in the olfactory bulb and change in frequency over time, resulting in a broader peak than the constant frequency line noise. Endogenous spiking activity recorded from the slice is shown in Figure 4.11. These data were bandpass filtered in MATLAB (200Hz$\leq f_{bp} \leq$6kHz) to remove LFP activity.

4.2.7 Conclusion

This paper has presented a highly scalable architecture and demonstration of a CMOS sensor array with metal electrodes and optical sensors. The design was based upon the characteristics of extracellularly recorded neural potentials. Due to a distributed ADC and a compact low-noise front-end amplifier, the system achieves a high-spatial resolution of 50μm without compromising recording

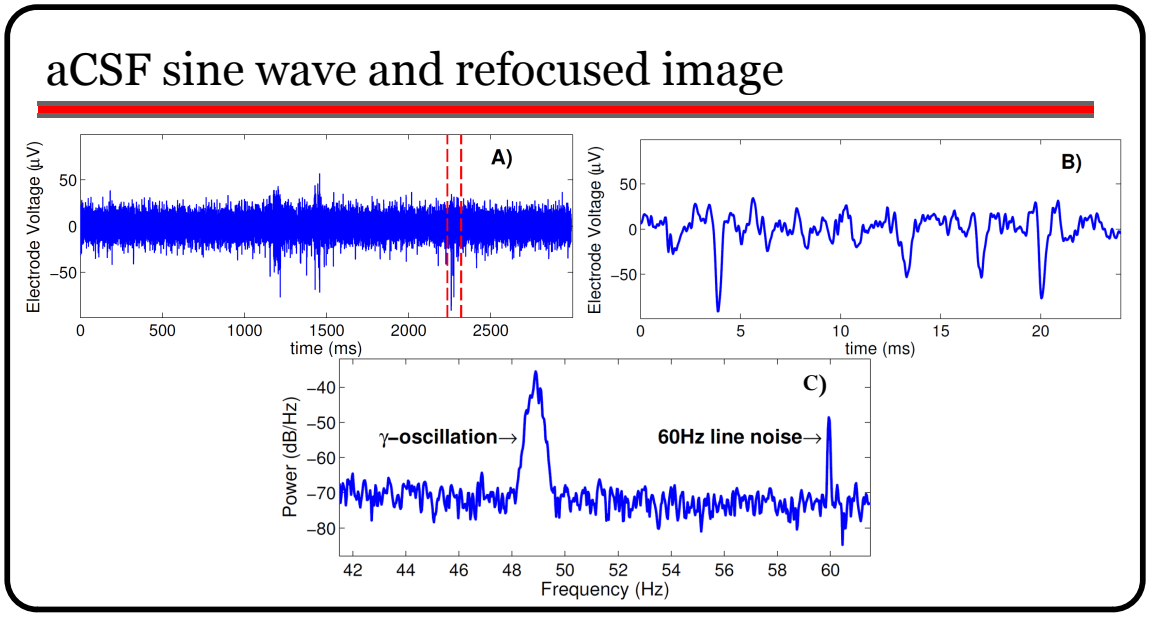

Figure 4.11: Recording of spiking activity and LFP gamma oscillation

area or sampling rate. The system requires simple post-processing to encapsulate the wirebonds and platinize the aluminum electrodes for biocompatibility. The system can perform lensless imaging to localize external electrodes by computationally refocusing information from ASPs. Action potential and LFP recordings from a mouse olfactory bulb slice were also presented. The presented system will enable unique neurophysiological experiments on a chip.

4.3 A 50um-pitch, 1120-Channel, 20kHz Frame Rate Microelectrode Array for Slice Recording

This section discusses the second-generation CMOS microelectrode array for direct recording of neural slices in 180nm CMOS (published in [45]). The array

consists of 1,120 active channels with 50μm pitch. Every sensor site has a frontend low-noise amplifier and photopixel for correlating optical stimulus with electrical activity. The frontend is AC-coupled and achieves area-efficiency by integrating the large input capacitor and recording electrode directly over the circuitry in conjunction with a single T-capacitor feedback network. Degraded PSRR (63dB) and CMRR (21dB) from the single feedback network are overcome by utilizing a virtual shared reference, improving rejection to 84dB and 66dB, respectively. Despite a small area, the frontend amplifier has an input-referred noise of 4.3μVrms with tunable high- and low-pass corners with very little variation from site-to-site. Experiments from a transgenic mouse olfactory bulb slice are shown.

4.3.1 System Description

The system was implemented in a standard 180nm 1P6M CMOS process with 4μm-thick top metal with an active area of 2.2 x 1.7 mm^2 (Fig. 4.12). Fig. 4.13 illustrates the overall architecture of the system, comprising 28 rows with 40 recording sites each. Each channel contains a low-noise amplifier, tunable low- and high-pass filters, and a buffer. A 6-bit, 40 count gray code generated in the digital core selects between electrodes in each row to activate each site's buffer. Each row includes a switched-capacitor programmable gain amplifier (PGA) with 4 gain settings (18 dB, 21.5 dB, 24 dB, or 26 dB), a unity gain buffer, and a 10-bit SAR ADC. The PGA implements correlated double sampling by amplifying the difference between the electrode signal and the frontend reference through the channel buffer. Correlated double sampling removes the variable offsets of every source follower in the row circuit, reduces the amount of output swing

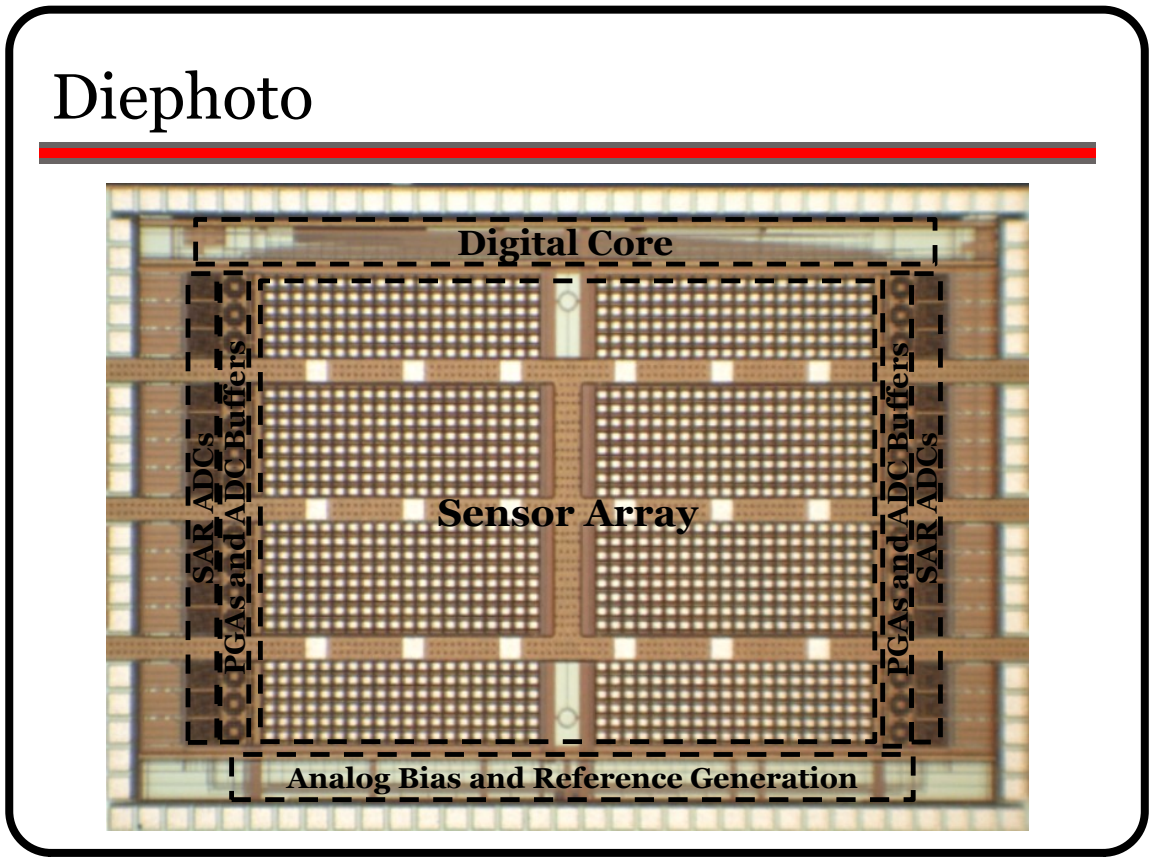

Figure 4.12: Die photograph of the sensor array

required by the PGA, and enables control of the PGA output DC level.

The dynamic range of extracellular neural signals from brain slices typically does not exceed 45dB, therefore the ADC uses 10 bit resolution to ensure that the quantization noise of the ADC is well below the noise floor of the recorded neural activity and frontend circuitry while providing margin for unexpectedly large input signals such as stimulation or perfusion artifacts. The SAR ADC uses a 5b/5b split capacitor array to reduce area and loading on the ADC buffer. Each ADC is operated at 800kS/s, effectively sampling each electrode at 20kHz and yielding a total aggregate data rate of 224Mb/s from 28 output pads.

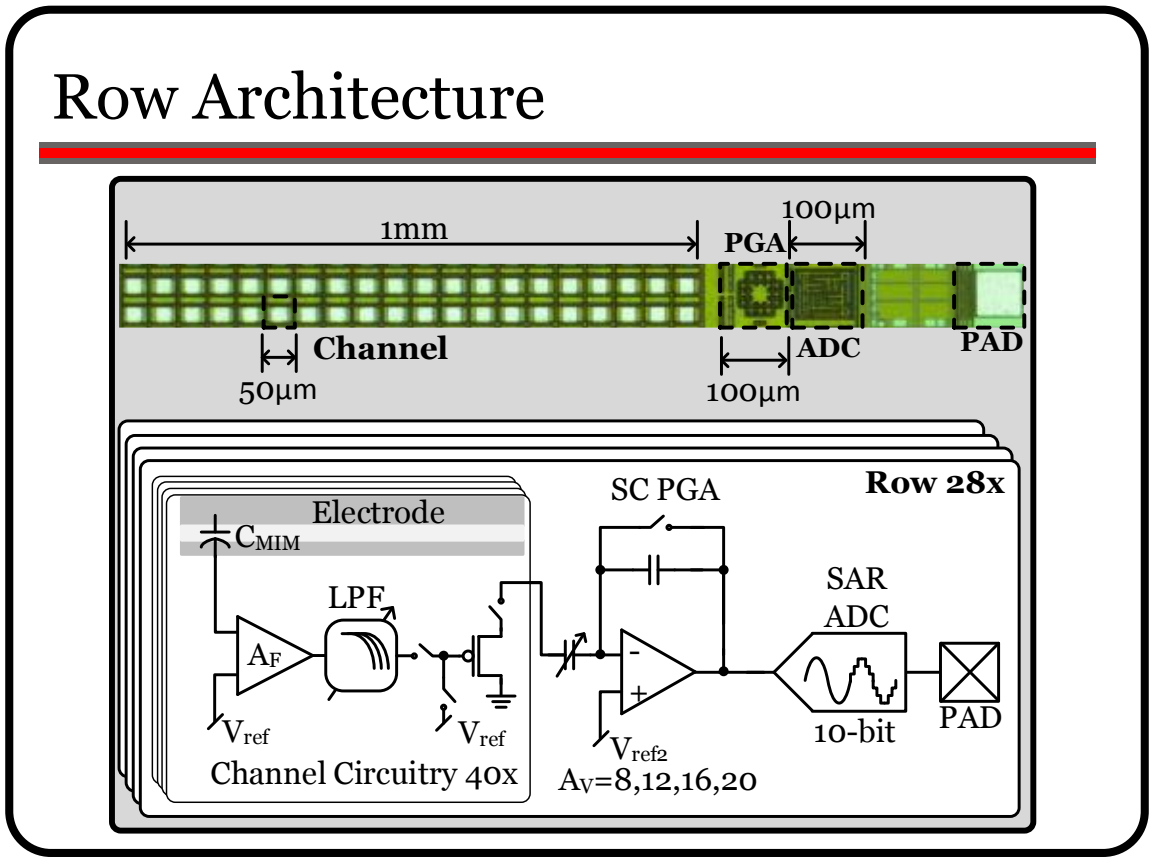

Figure 4.13: Block diagram of the row architecture

4.3.2 Unit-Sensor Design

A key challenge for high-density, robust recording is achieving low-noise and high gain with area-efficient circuits [21]. Local filtering and amplification are necessary at each electrode to limit the thermal noise from the electrode interface (typically the dominant noise source for MEAs) and prevent aliasing under rapid multiplexing of signals. High gain and low output impedance at the pixel-level are desired to relax noise requirements of subsequent signal conditioning stages and to reduce crosstalk and EMI pickup [61]. Traditional neural amplifiers use capacitive coupling to block low frequency electrochemical offsets and

capacitive feedback to accurately set the gain across channels [7]. The gain is set by the ratio of the input capacitor to the feedback capacitor. The feedback capacitor value is typically set by parasitics and matching (roughly 150fF), resulting in very large areas for sufficient gain (40dB). Therefore, this work uses a capacitive T-network to decouple the gain from the maximum capacitor ratio, allowing a closed-loop gain of 41.2dB with ratios less than 17:1 (Fig. 4.14). The T-network creates an effective gain feedback capacitance (C_{gain}) of

$$C_{gain} = \frac{C_2 \cdot C_4}{C_2 + C_3 + C_4} \tag{4.2}$$

where the closed loop gain of the amplifier is C_1 / C_{gain}. C_{gain} is now 22.5fF, which is smaller than the minimum size MIM capacitor without sacrificing matching.

While decreasing the size of the input capacitor can save significant area, it has a few caveats. Ensuring C_1 is much larger than the feedback network and the input capacitance of the amplifier ($C_{in,amp}$) mitigates most of the issues. The input-referred noise of the open-loop amplifier ($V^2_{n,amp}$) sees a different transfer function than the input signal, and is related to the input of the system as

$$V^2_{n,fdbk} = \left(\frac{C_1 + C_{ff} + C_{in,amp}}{C_1}\right)^2 \cdot V^2_{n,amp}, \tag{4.3}$$

where

$$C_{ff} = \frac{C_2 \cdot (C_3 + C_4)}{C_2 + C_3 + C_4}. \tag{4.4}$$

C_{ff} is the feed-forward capacitance of the T-network as seen from the input of

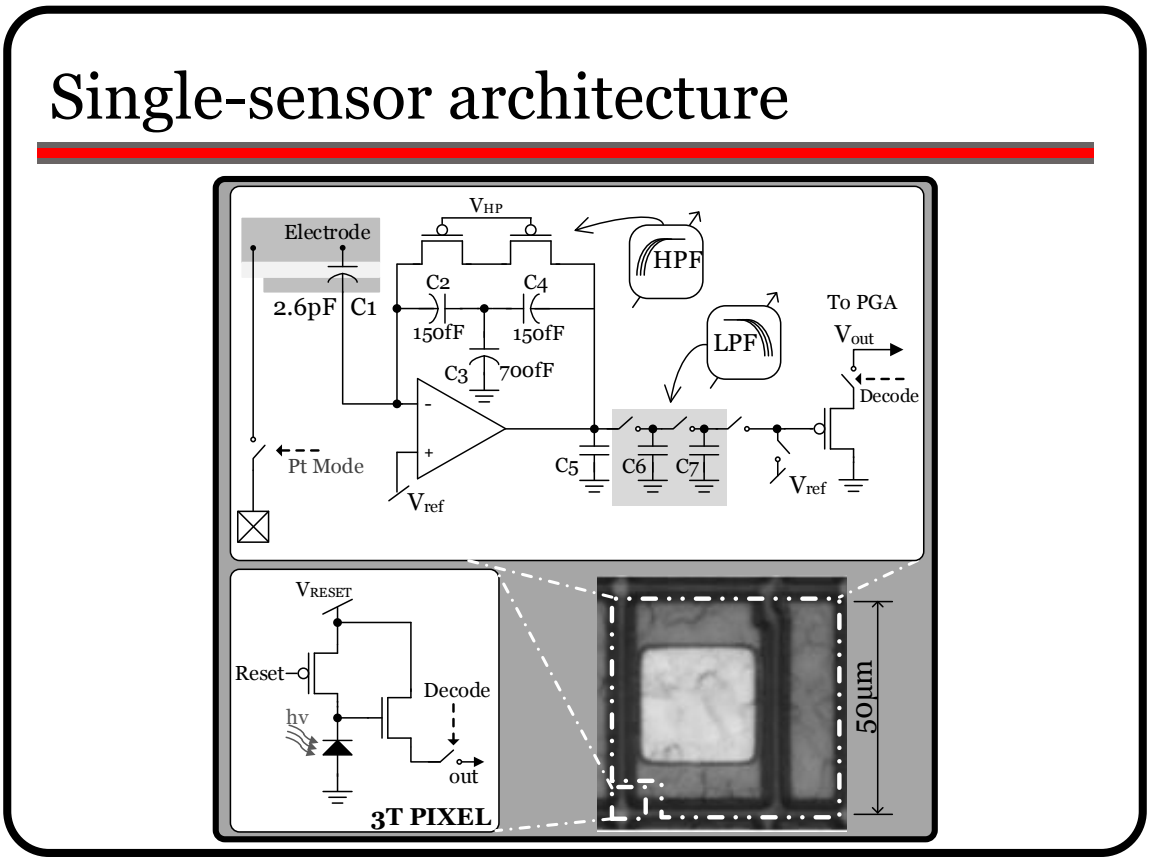

Figure 4.14: Unit sensor with frontend amplifier and photopixel

the amplifier. The input pair is large to suppress flicker noise; however, increasing its area increases the noise gain.

Note that equation 4.3 neglects kT/C noise from the feedback network, which is dominant at low frequencies due to a small C_{gain}. Once again, a large input capacitor (high gain) will mitigate noise injected by the feedback impedance. kT/C noise can also be shifted by the high-pass corner (ω_{HP} = 1 / $R_{eff}C_{gain}$). The high-pass corner is controlled by a PFET in triode with a large incremental resistance (R_{eff}). The main purpose of the high-pass corner is to keep low frequency artifacts from saturating the system. The low-pass corner

of the amplifier is also adjustable to selectively filter for LFP or spikes, optimizing the array for different experimental conditions. The low-pass corner is set by a switched-capacitor filter on the output of the amplifier ($\omega_{LP} \approx f_{sw}C_6$ / C_7). Since the corner is determined by a ratio of local capacitors and a switching frequency, the corner is robust against mismatch.

Fig. 4.15 shows the transistor-level schematic of the amplifier. The amplifier uses a low-power folded-cascode topology with source degeneration resistors to reduce noise and improve matching [87]. Since there is no matching feedback network for the reference, the amplifier is more susceptible to power supply fluctuations coupling through the C_{GS} of M_1. We use principal component analysis (PCA) to remove common-mode signals, improving PSRR. PCA is effective because the array records from several sites simultaneously distributed across a large spatial area, meaning that the recorded neural signals are not likely to be correlated. Noise from the differential pair biasing cannot be removed through PCA, so R_1 was inserted to degenerate noise from M_{b1}.

The electrode circuitry was designed to interface directly to neural tissue in an aqueous environment. The interface electrode is formed by a 25μm by 25μm passivation opening over the top plate of the input MIM capacitor (Fig. 4.15). M4 was used exclusively as a ground plane in the sensor area to provide additional electrical, ionic and light shielding. One corner is unshielded to allow light to pass to the photodiode. A switch connects the top plate of the input MIM capacitor to an externally controlled voltage to enable the controlled platinization of electrodes.

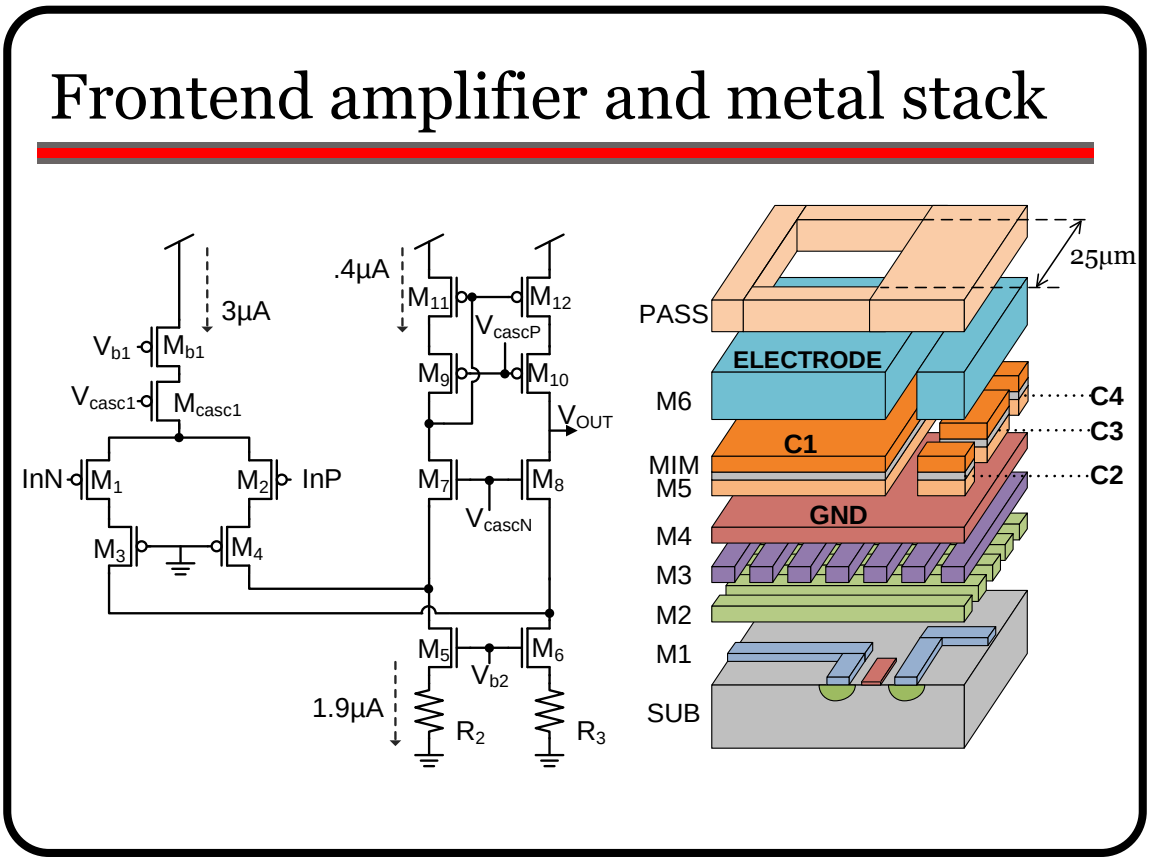

Figure 4.15: Frontend amplifier schematic and electrode metal stack

4.3.3 Measurement Results

While significant variation in high-pass corner frequencies was expected across the chip, direct measurements shown in Fig. 4.16 indicate that the variation (σ = 0.8Hz, roughly 10%) is sufficiently small for biopotential applications. Variation in the midband gain (with gain loss from the buffer) and the low-pass corner were measured to be 0.6% and 1.2%, respectively, demonstrating good matching across the array. The frontend input-referred noise was 4.3μVrms (measured from 20Hz - 50kHz) from a 1.8V, 3.8μA supply, corresponding to a noise efficiency factor (NEF) of about 3.4 for a bandwidth of 20Hz to 9kHz, com-

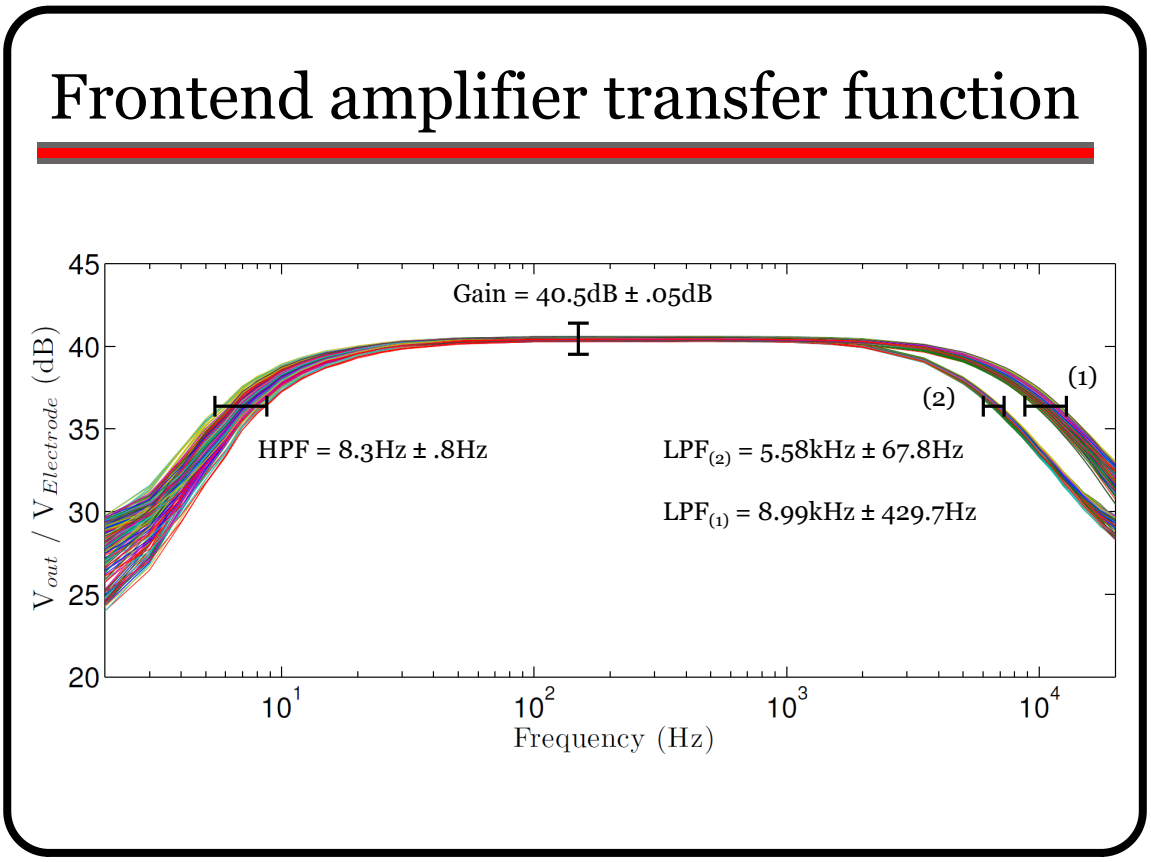

Figure 4.16: Frontend amplifier transfer function

petitive with state-of-the-art neural preamplifiers with much larger areas [7]. The measured average PSRR and CMRR across the array was 63dB and 21dB, respectively. After performing PCA, the average PSRR and CMRR were improved to 84dB and 66dB. The backend signal conditioning achieved an ENOB of 8.2 bits at 800kS/s, demonstrating a sufficient SNR for neural recording. The power consumption of all analog frontends, PGAs, and ADCs was 14.1mW, corresponding to 12.6μW per channel.

Post-processing the system for neural recordings involved defining a well around the sensor array with silicone, and then encapsulating the bondwires

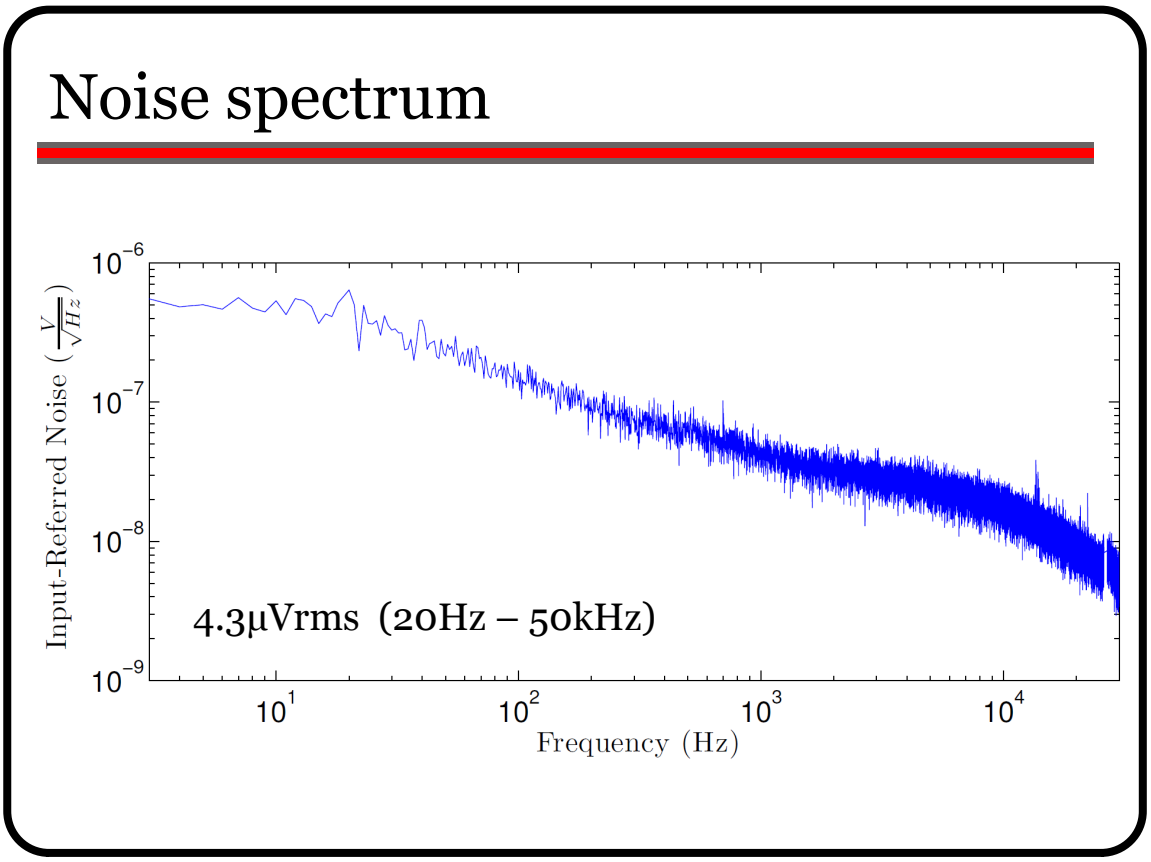

Figure 4.17: Measured frontend noise spectrum

with epoxy. Since the exposed metal is aluminum, which corrodes easily in saline and is cytotoxic, the electrodes were electroplated with platinum black. Platinum is nontoxic and decreases electrode impedance, hence reducing the noise contributions from the electrode interface. Electroplating was performed by filling the well with platinizing solution, applying a positive potential to a platinum counter-electrode, and holding the electrodes at a fixed potential by activating internal platinization circuitry.

The functionality of the system was verified by recording spiking and LFP activity from a 300μm-thick mouse olfactory bulb slice. The tissue was hori-

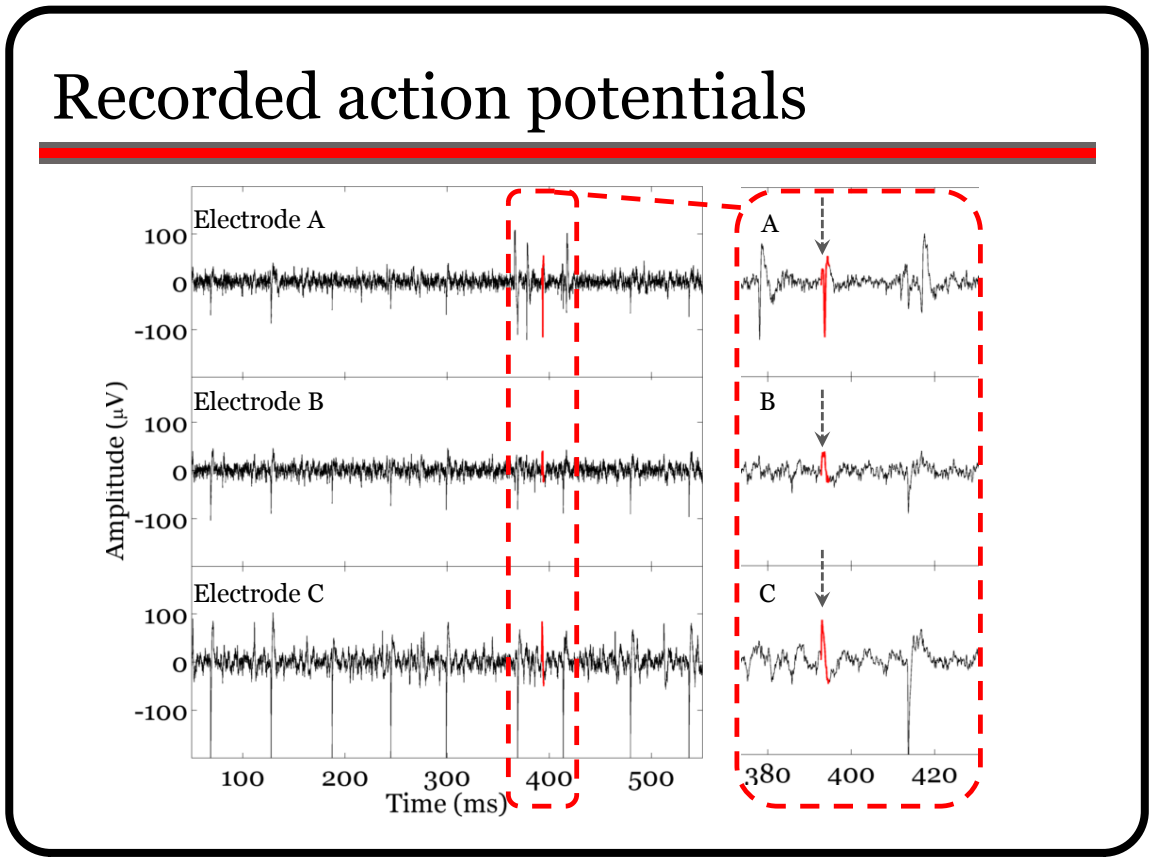

Figure 4.18: Recorded action potentials on multiple electrodes

zontally sliced with a vibrating microtome and then placed in oxygenated 34C aCSF. Fig. 4.18 shows recording of endogenous neural activity recorded from three electrodes. Due to the fine pitch of the array, single action potentials can be spatially oversampled.

An iPad with Retina display approximately 1m above the array was used to generate test inputs for the photopixels. A 20mm, f1.8 Sigma lens was used to focus images from the screen onto the sensor array. Fig. 4.19 shows that the array can be used as feedback to focus optical stimulus with close to cellular precision. Furthermore, the array can take video to capture time-varying stimuli.

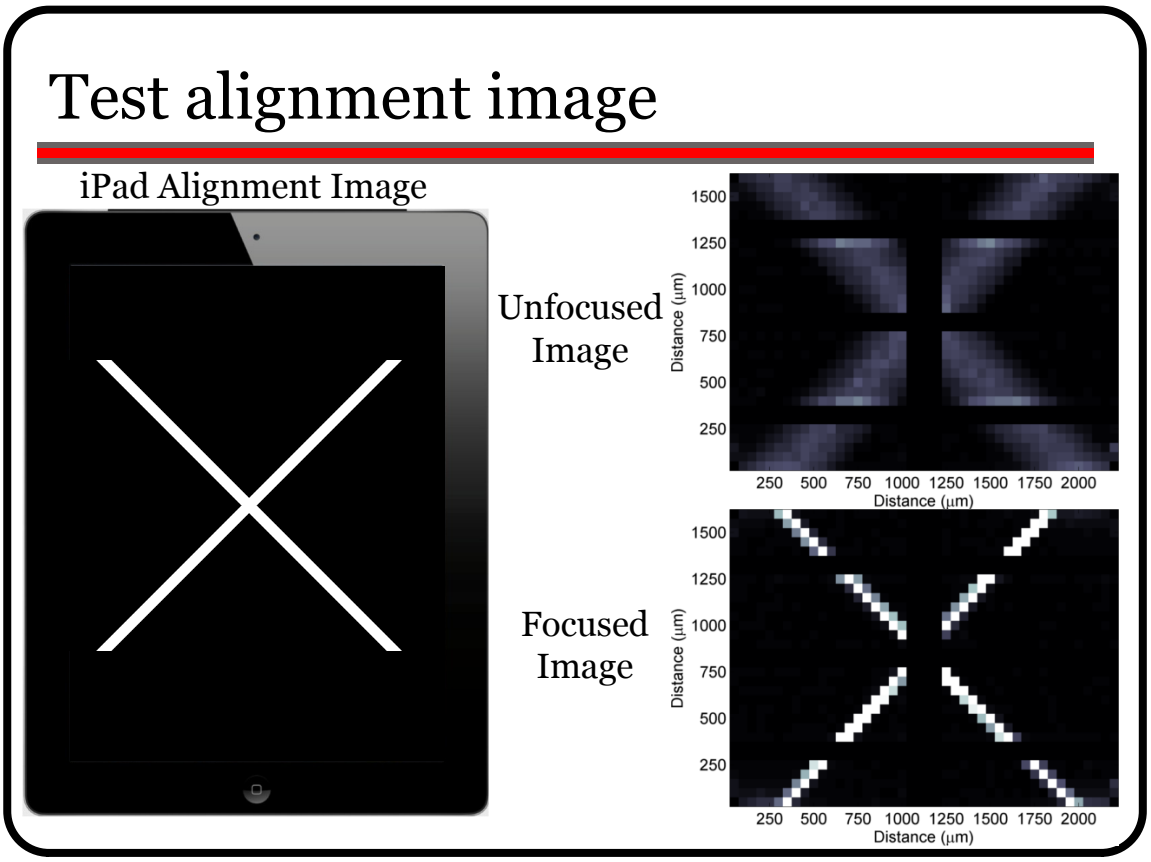

Figure 4.19: Test alignment images captured by array

Fig. 4.20 shows example still-frame images captured by the sensor array

demonstrates that the sensor array can be used to determine whether optical stimuli are focused and capturing video.

4.3.4 Conclusion

This work presented a 1,120-channel electrode array for neural slice recording. Despite a small area (50μm by 50μm), each channel had a power-efficient,

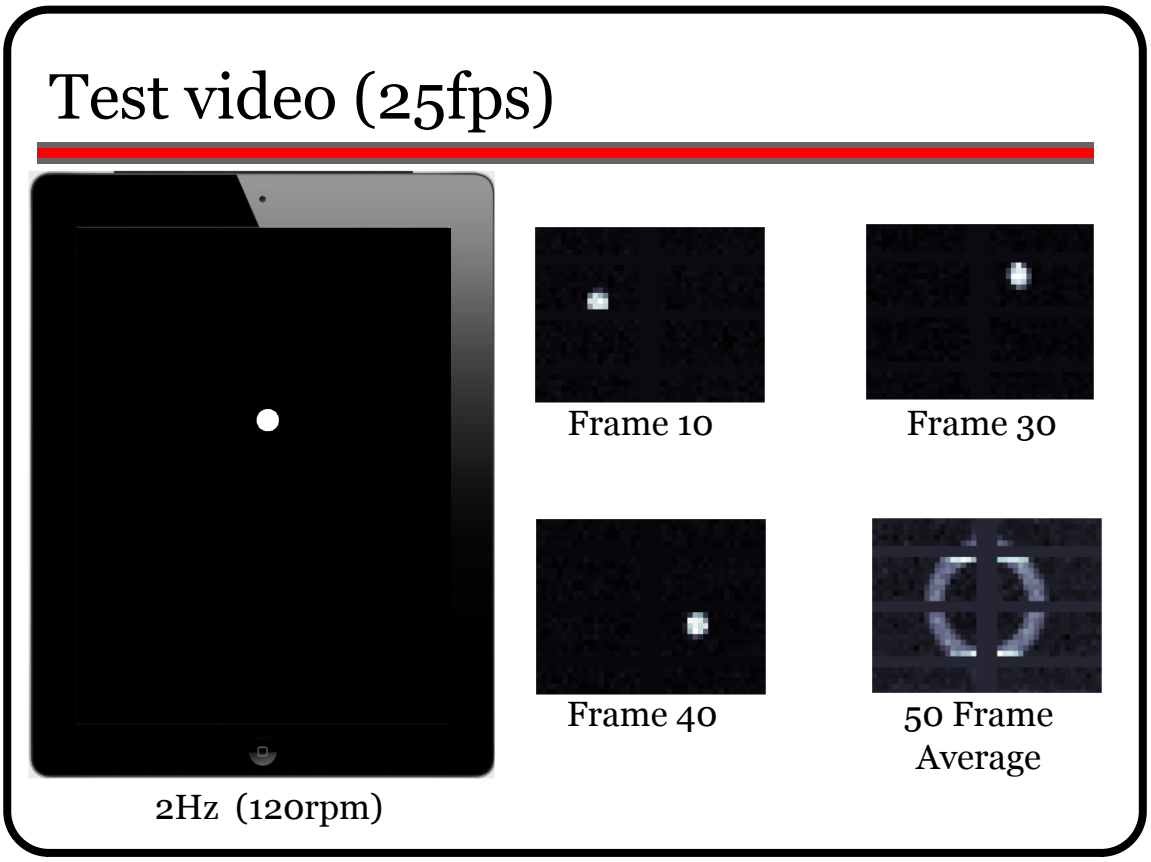

Figure 4.20: Test video captured by array

low-noise amplifier (4.3μVrms) and photopixel for optical sensing. The array demonstrated good matching across channels and was used to record neural activity from a neural slice.

CHAPTER 5

EFFICIENT IMAGING: A HIGH-SPEED POLAR-SYMMETRIC IMAGER FOR REAL-TIME CALIBRATION OF ROTATIONAL INERTIAL SENSORS

This sections presents a high-speed (> 1kfps), circular, CMOS imaging array for contact-less, optical measurement of rotating inertial sensors (published in [43]). The imager is designed for real-time optical readout and calibration of a MEMS accelerometer revolving at greater than 1000rpm. The imager uses a uniform circular arrangement of pixels to enable rapid imaging of rotational objects. Furthermore, each photodiode itself is circular to maintain uniform response throughout the entire revolution. Combining a high frame rate and a uniform response to motion, the imager can achieve sub-pixel resolution (25nm) of the displacement of microscale features. In order to avoid fixed pattern noise arising from non-uniform routing within the array we implemented a new global shutter technique that is insensitive to parasitic capacitance. To ease integration with various MEMS platforms, the system has SPI control, on-chip bias generation, sub-array imaging, and digital data read-out.

5.1 Introduction

MEMS accelerometers are a favorable alternative to GPS for autonomous navigation in GPS-denied environments. However, bias and scale-factor drift remain major obstacles for precise and long-term position tracking using MEMS accelerometers [83]. One possible solution to these time-varying, non-deterministic errors is to calibrate the inertial sensor in real-time by measuring the displacement of a revolving accelerometer in two opposite phases. Comparing these readings differentially means that the signal from the applied force

should add while the drift and offset components cancel.

Contact-based read-out from a revolving sensor is problematic, favoring contactless (optical) readout. This can be implemented by probing the accelerometer with an overhead illumination and imaging the generated diffraction pattern (Fig. 5.1). Ideally this system is contained within a small volume (<1000mm^3) and all the processing for calibration is performed on-chip. Such a setup requires an image sensor to track a quickly revolving diffraction pattern and sense nanometer shifts within the pattern. Traditional imagers record on a rectangular grid which is not well-suited for efficiently imaging rotating objects. The number of pixels on a rectangular grid required for tracking the object is proportional to D^2, where D is the diameter of the circular path. Since the relevant features are along a particular circular contour of the inertial sensor, many interior and corner pixels are superfluous. This both limits the maximum frame rate and increases the output data rate. Another difficulty with using traditional imagers for real-time calibration is that the computation of angle from Cartesian coordinates requires the implementation of an inverse tangent function. This processing overhead is significant considering it must be done in real-time and on-chip.

Therefore, this work presents an ASIC imaging sensor that efficiently captures the diffraction pattern generated by a rotating inertial sensor as part of a low-power calibration system. For more efficient spatial sampling and simpler processing, this work uses an array of pixels uniformly distributed along polar coordinates to directly extract the angular position.

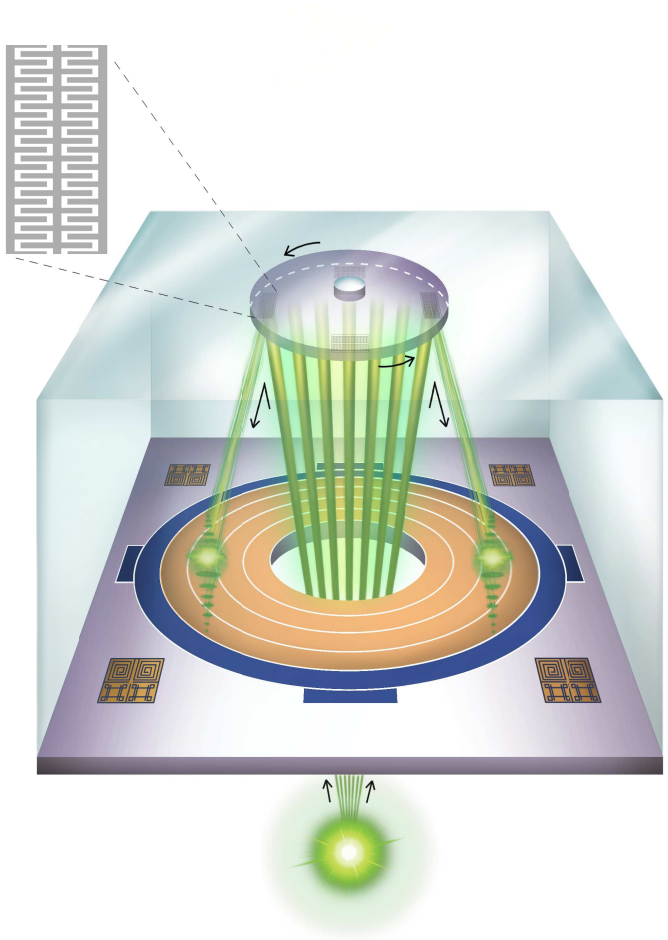

Figure 5.1: System for continuous accelerometer calibration

5.2 Theory

A circular geometry reduces the pixel count to be proportional to the circumference of the rotation path which, scales linearly with diameter. For instance, imaging the circumference of a 1mm diamater ring with 5μm-pitch pixels with an 8-pixel wide circular array requires only 1000 pixels, while a rectangular array would require 40,000 pixels.

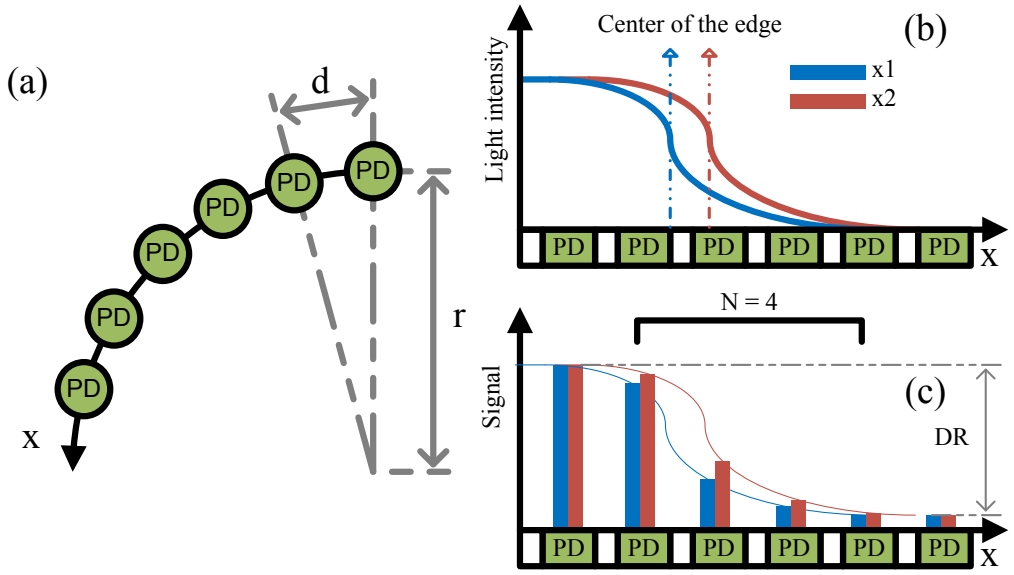

Figure 5.2: Concept of linear interpolation with uniform circular response

Since the pixels are uniformly distributed in a circular array, pixel coordinates directly provide a measure of angular position. Therefore the angular location of an imaged spot distributed across several pixels can be measured to within $\frac{2\pi}{N}$ radians by simply considering the position of the center (brightest) pixel, where N is the number of pixels along the circumference. A much more precise estimate of angle can be computed by considering the outputs of all the pixels that span the imaged spot and interpolating. Fig. 5.2 illustrates the process of achieving sub-pixel resolution of edges as they revolve around the array. Ideally the edge can be computed with an arbitrary precision far less than the pixel size; however, the resolution will be dictated by the noise of the system (σ_{system}, shot noise from the photodiodes and noise from the readout circuits) and its dynamic range (DR). The spatial resolution (σ_x) is then given by:

$$\sigma_x = \frac{d}{\sqrt{n}} \cdot \frac{\sigma_{system}}{DR} = \frac{d}{\sqrt{n} \cdot SNR} \tag{5.1}$$

where d is the pitch of the pixels and n is the number of pixels that image the edge. Given an imager radius r, spatial resolution can be converted to an angular resolution by $\sigma_\theta = \frac{\sigma_x}{r}$.

The precise computation of angle also requires the response of individual pixels to be independent of location in the array. This is achieved by using disc shaped photodiodes for each pixel. Additionally, photodiode shot noise is shape-dependent and is reduced by using a circular diode. This is due to an evenly distributed electric field which reduces stress-induced leakage current [75], [55].

5.3 System Design

5.3.1 Architecture

Fig. 5.3 shows the top-level functional diagram of the proposed system. The sensor array is organized into 4 concentric circular bands of pixels. Each band is 8 pixels wide with the circumferential pixel count increasing from 320 pixels for the innermost ring (R0) to 512 pixels for the outermost ring (R3). The region inside the innermost ring of photo-detectors can be hollowed out to facilitate on-axis illumination of the rotating intertial sensor.

The top-level control circuitry includes the addressing and timing controls, serial programming interface(SPI), and bias, reference, and supply voltage generation. Addressing logic selects one of four pixel bands for readout. Within each band the pixels are addressed on a polar coordinate system via angular

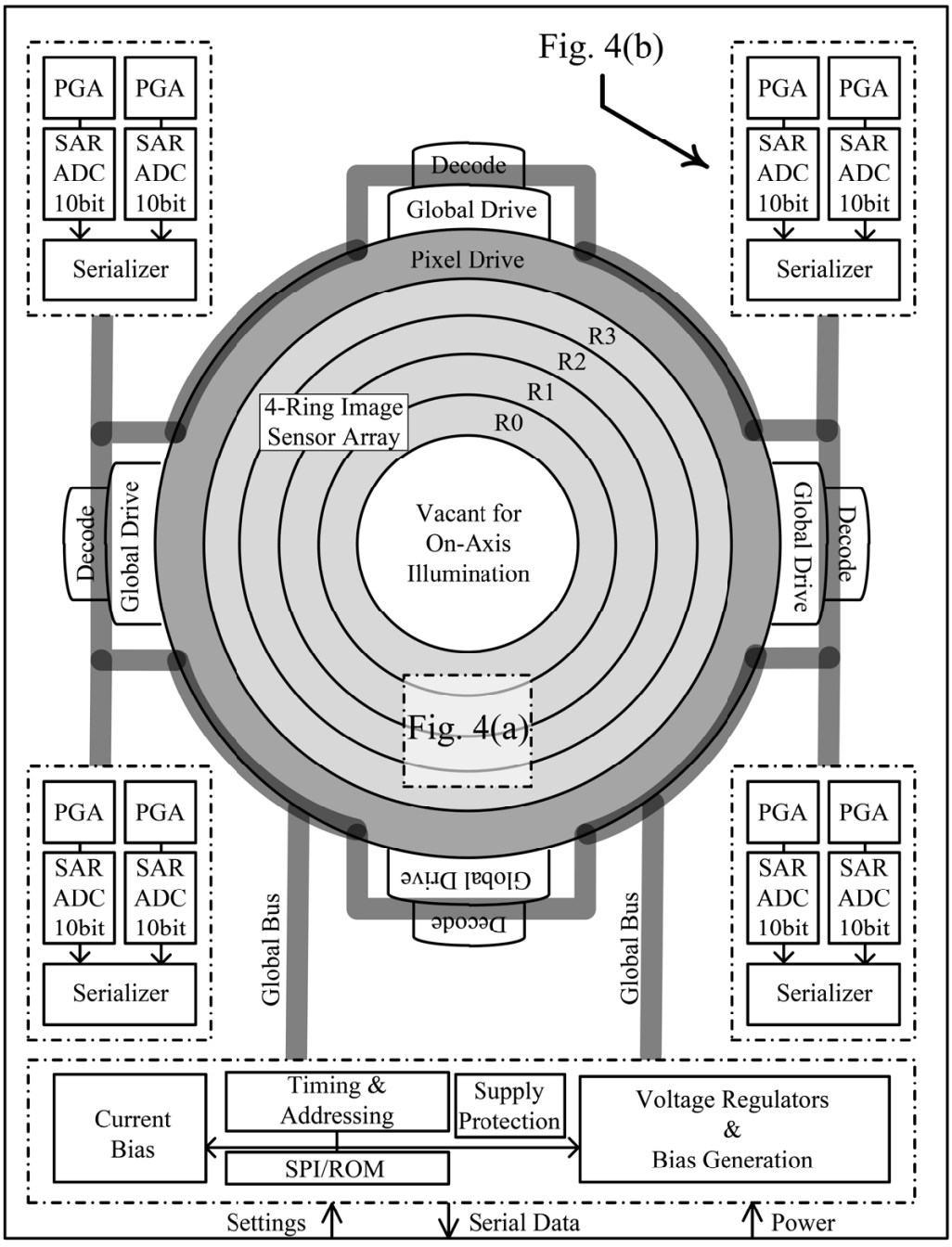

Figure 5.3: Top-level architecture of sensor system

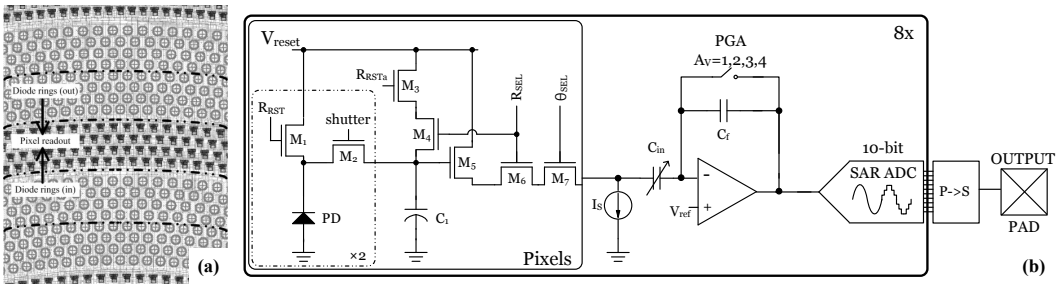

Figure 5.4: Circular photodiode layout and signal pathway block diagram

and radial selection controls. The address controls are designed to permit highly flexible readout from any desired sector of the array. By varying the range of valid addresses operation can be limited to a particular angular window. This mode of operation is intended to do allow for extremely high frame rate operation while reducing data-handling requirements. The position and extent of the angular window can be programmed in real time via the SPI to permit object tracking. Bias and reference levels are generated on-chip to reduce the number of I/O pads and off-chip components required.

A custom place and route program was implemented to efficiently layout the circular rings, as descibed in Section *B*. Given the circular symmetry of the system, the array was assembled as 4 identical quadrants each with dedicated quad-level back-end circuits and shared top-level circuits. Every pixel output is amplified by a programmable gain amplifier (PGA) and is digitized by a SAR ADC. The ADC outputs are time-interleaved by a serializer and output on a single pad (Section *C*).

5.3.2 Circular Layout

Since the sensor array is intended to work with various MEMS platforms with ongoing revisions, the array design process had to be flexible to cater to rapid redesign. Custom CAD automation, similar to [89], was required for polar coordinate place and route of pixels and readout, since both design tools and CMOS fabrication process are inherently optimized for Cartesian coordinate design. We developed a custom software tool to generate circular photodiodes for various radii and to automate the placing of photodiodes, pixel readouts, image core drivers, and decoders. Furthermore, the custom placement software placed pins to guide the Cadence Virtuoso Chip Assembly Router (VCAR).

Fig. 5.4(a) shows the completed layout of two pixel rings and a row of readout circuits. The CMOS process used for the proposed sensor array has 6 metal layers, where the first two layers were used for local routing (photodiode to pixel readout), metal 3 for optical shielding to avoid optical mismatch between pixels, while layers 4 and 5 were used for global routing (addressing and global shutter operations). The top metal layer (metal 6) was not used for imaging core layout due to thickness and minimum width constraints. In order to maximize fill-factor, we separated photo-diode and pixel readout circuits into separate bands to share the readout circuits by two neighboring bands.

5.3.3 Pixel Readout Circuitry

Fig. 5.4(b) shows the signal pathway from the photodiode to the output. In order to limit motion artifacts, the sensor uses a global shutter scheme for photodetection. In this approach, all photopixels are reset via transistor M1 simul-

taneously at the start of the frame. Photocurrent is integrated during the frame period after which the accumulated charge is transferred to a storage capacitor (C1) by turning on a shutter transistor (M2). Also note that the shutter transistor also selects the diode ring connected to the readout circuitry. The capacitor, implemented with an NFET, forms a part of the pixel readout circuitry and is placed in a separate band from the diodes. Routing from photodiode to the readout circuitry varies from pixel to pixel, resulting in fixed pattern noise from the mismatch between the diode capacitance and the storage capacitor ratio. To address this problem we implemented a second global shutter scheme that is insensitive to the capacitance on the photodiode node. In this mode, the reset switch (M1) is unused. We hold the transfer switch (M2) at a voltage $V_{shutter}$<VDD. This charges up the diode to a reset voltage equal to $V_{shutter}$-V_T. The charge required to restore the diode to the reset level is supplied by the storage capacitor. This simultaneously reads the photo-charge onto the storage node and resets the photodiode.

5.3.4 Backend Readout Circuitry

Subsequently the storage cap is read out through a source follower (M5). To remove offsets and low frequency noise from the source follower, the array implements correlated-double sampling with the switched-capacitor PGA. On the first phase, the pixel signal is sampled onto the input capacitor and then on the second phase the feedback switch is opened and the reset level is sampled onto the input capacitor, resulting in an output voltage of $\frac{C_{in}}{C_f}$·(V_{pixel} - V_{reset}) + V_{ref}. The amplified signal is then digitized by a 10-bit SAR ADC. Note that the DC output level of the PGA, set by V_{ref}, is adjusted to be close to the high reference

of the ADC to maximize the dynamic range since $V_{pixel} \leq V_{reset}$. The ADC uses a 5b/5b split capacitor array to reduce area and input capacitance. Outputs from all eight ADCs are serialized and then transmitted off-chip through a single pad.

5.4 Measurement Results

We implemented the image sensor in a 180nm CMOS process. A die photo of the implemented system is shown in Fig. 5.5, where (1) is the PGA and ADC, (2) is the space for on-axis illumination, (3) is the concentric pixel array rings and (4) is the digital control and bias generation. The inset is the layout of read-out circuitry and circular photodiodes. Each photodiode is 7μm in diameter. A single band is 80μm-wide with the inner diameter ranging from 1.08mm for the innermost ring to 1.8mm for the outermost ring. As can be seen in the inset of Fig. 2.10, alternating rings of diodes were angularly staggered to achieve a closely packed arrangement. The fill factor calculated from the layout is 33.5%. The central disc within the array core was hollowed out using an LPKF PCB prototyping laser. A 100μm clearance width was used during the drilling process to avoid damage to the imaging array. LED back-illumination can be observed through the cavity in the die photo of Fig. 5.5.

The chip test setup used a 3.45V supply for the on-board regulators, an external current reference for bias generation, and clock signals for the digital addressing controls. At 2MHz clock frequency, corresponding to 16fps for full array readout, the chip consumed 15.5mW. Digitized images were acquired in 8 channel parallel readout mode.

An iPad with Retina display with a maximum brightness of 455cd/m^2 ap-

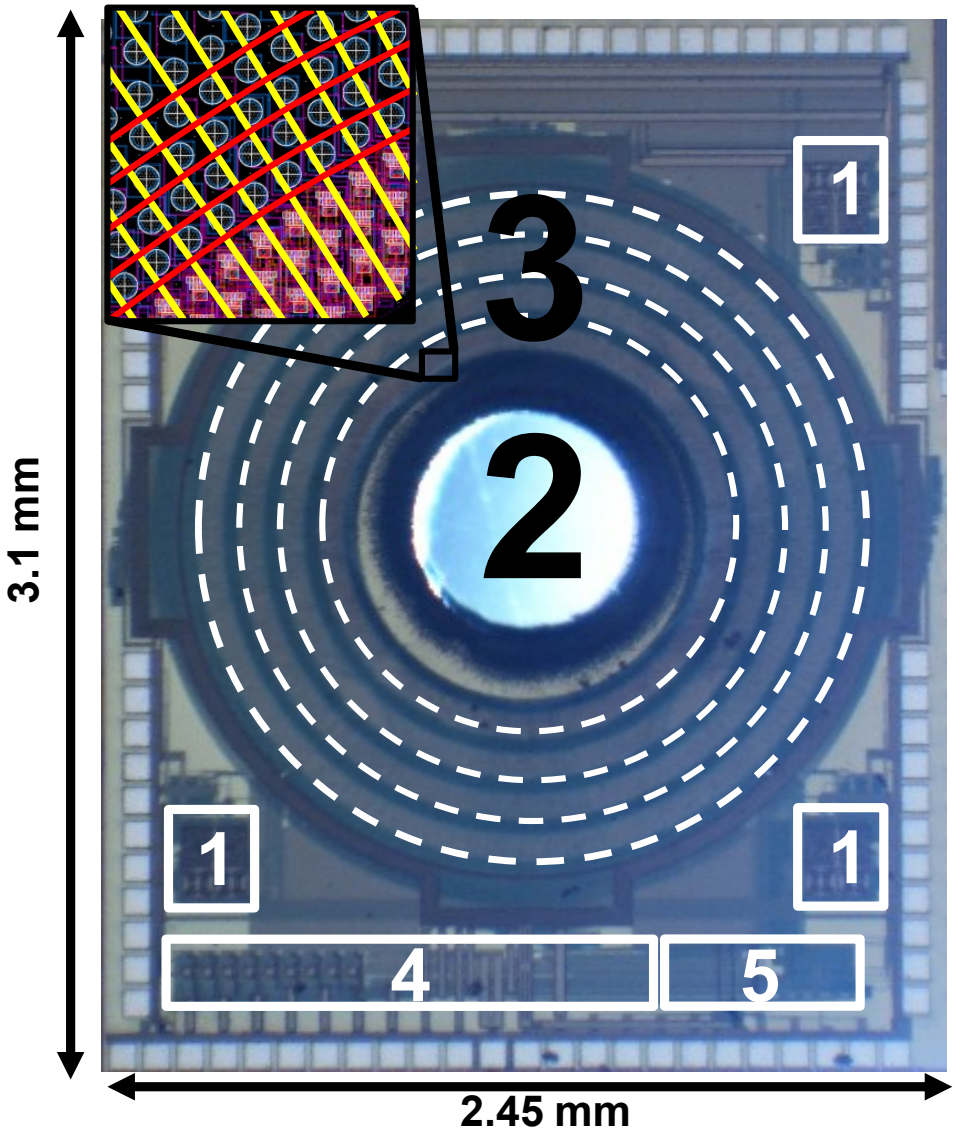

Figure 5.5: Die photo of implemented system

proximately 1m above the array was used to generate test inputs for the sensor array. A 20mm, f1.8 Sigma lens was used to focus images from the screen on to the array. A magnification of 30:1 was used to project the screen on to the sensor's active area, meaning 1 pixel on the display corresponded to 3μm on the sensor. Prior to testing, the display and optics were aligned to the sensor by centering a cross-hair alignment mark. Fig. 5.6(b) shows the captured responses

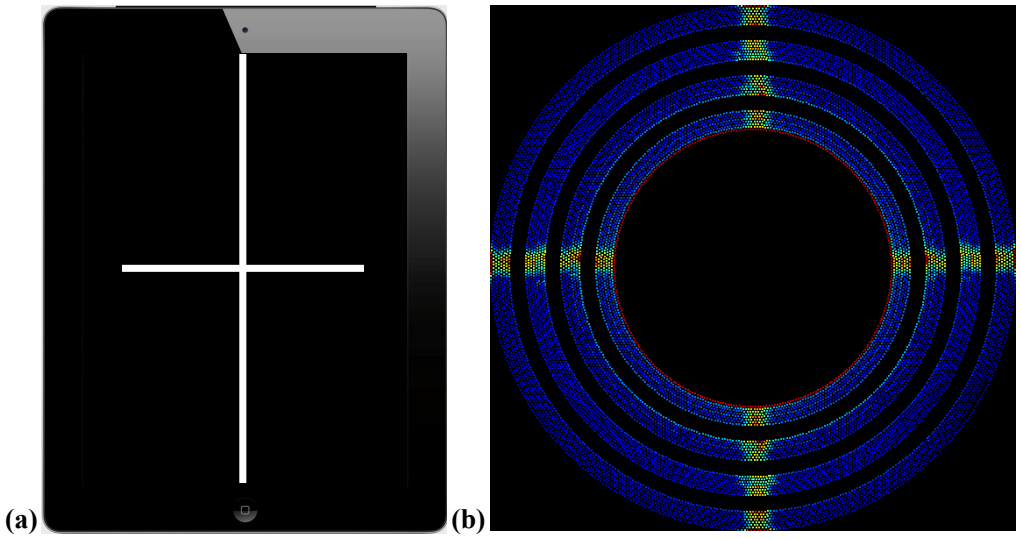

Figure 5.6: Alignment image on iPad and captured alignment image

of all four array sub-bands to the alignment mark. The circular distribution of pixels reduces the pixel count required to image a 1.5mm diameter path by a factor of 20 compared to a rectangular grid.

To measure the angular precision of the sensor, we imaged two fixed illumination spots, one of which had a slightly displaced edge. For each test image 55 consecutive frames were acquired. Fig. 5.7(a) shows the test image with the shifted edge highlighted. The spot spanned a 240μm arc on the sensor covering a 7x20 band of pixels. The measured pixel outputs along a single arc, shown in Fig. 5.7(b), show a linear pixel shift along the angular coordinate corresponding to the angle shift in the input. The weighted average of the angular coordinates of the illuminated pixels was computed to determine the precise position of the illumination. The calculated angle has a standard deviation of 32μrad across the 55 test frames. This translates to a spatial precision of 25nm for a ring radius of 780μm.

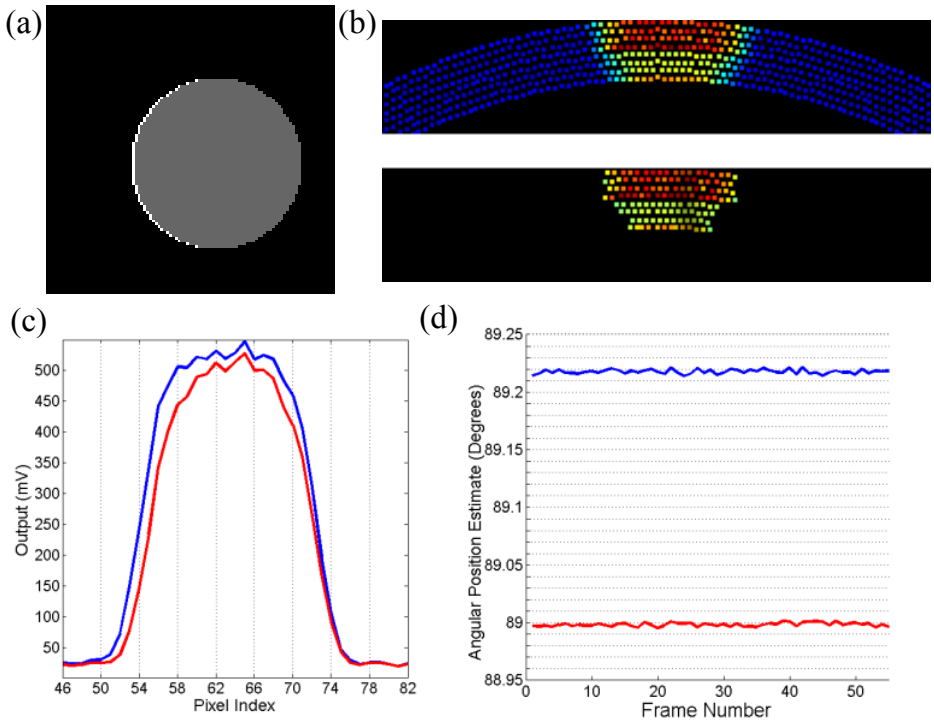

Figure 5.7: Angular position of spot after edge shift ($\sigma_\theta = 32\mu rad$).

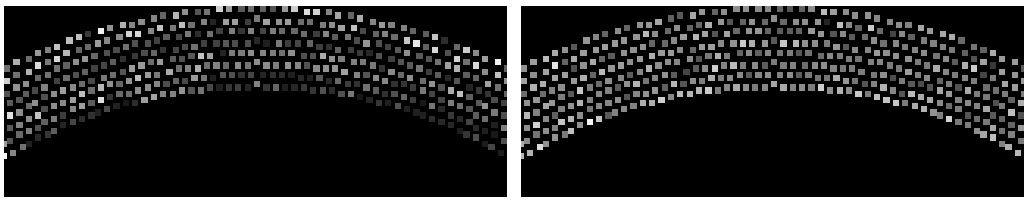

Figure 5.8: Conventional (left) and capacitance-insensitive (right) reset.

A comparison of low light level images taken using conventional global shutter and our diode capacitance insensitive global shutter is shown in Fig. 5.8. Fixed pattern noise was reduced by a factor of two with the capacitance insensitive global shutter enabled.

5.5 Conclusion

We demonstrated a new sensor architecture that uses photopixels uniformly distributed along polar coordinates to directly extract angular position. We measured a sub-pixel angular resolution of 32μrad, corresponding to a 25nm spatial resolution. The system also utilized a novel global shutter technique that reduced fixed pattern noise by a factor of two. In summary, this work presented a fully-integrated CMOS imaging array for optical rotation measurement of MEMS intertial sensors.

APPENDIX A

APPENDIX OF CHAPTER 1

To find the effect of crosstalk in the stack, we start by looking at the y^{th} layer of the input stack, which has 2^{y-1} differential pairs. We can describe the tail currents of these differential pairs with a vector $\vec{I}_y$ of length 2^{y-1}. These currents include a bias current $\vec{I}_y/2^{y-1}$ as well as contributions from each of the inputs above it in the stack. In general, each input voltage V_x ($x < y$) contributes to the vector of tail currents as

$$\vec{I}_{xy} = \frac{g_m}{2^{y-1}} \cdot \vec{k}_{xy} \cdot V_x \tag{A.1}$$

where g_m is the composite transconductance common to each combined set of differential pairs and the vector $\vec{k}_{xy}$ is also of length 2^{y-1}, consisting of alternating strings of 1 and -1 of length 2^{y-x-1}. For example, if $y = 4$ and $x = 2$, $\vec{k}_{xy}$ would be a column vector with entries: 1,1,-1,-1,1,1,-1,-1. Note that for a given value of y, each vector $\vec{k}_{xy}$ is orthogonal to those with different values of x. The 2^{y-1} differential pairs then steer these currents, due to mismatch, with coefficients $\vec{\Delta}_{VTHy}/(2\eta V_T)$, where $\vec{\Delta}_{VTHy}$ is the vector of threshold voltage mismatches in each of the differential pairs in layer y, and ηV_T is the subthreshold exponent of the differential pairs. Once the output currents of these differential pairs are recombined in the output to extract the current due to V_y (equal to $g_m V_y$) the crosstalk current due to V_x will be

$$\vec{I}_{xy} = \frac{g_m V_x}{2^{y-1}\eta V_T} \cdot \vec{k}_{xy} \cdot \vec{\Delta}_{VTHy} = \alpha_{xy} g_m V_x \tag{A.2}$$

The entries in $\vec{\Delta}_{VTHy}$ are assumed to be Gaussian random variables with zero mean, and a variance of

$$\sigma^2_{VTHy} = \sigma^2_{VTH1} 2^{y-1} \tag{A.3}$$

where σ^2_{VTH1} is the variance in the mismatch of the top differential pair. The differential pairs lower in the stack are smaller in area by a factor of 2^{y-1} and have a proportionally larger threshold voltage variance. The crosstalk term, α_{xy}, then is zero mean and has a variance of

$$\sigma^2_{\alpha} = \left(\frac{1}{2^{y-1}\eta V_T}\right)^2 \cdot \sum_{i=1}^{2^{y-1}} k^2_{xyi} \sigma^2_{VTHy} = \sigma^2_{VTH1} \left(\frac{1}{\eta V_T}\right)^2 \tag{A.4}$$

which we note is independent of the values of x and y. The only subsequent circuits which can influence crosstalk are the recombination current mirrors, each of which distributes the ith current to the yth output with weight

$$I_{outiy} = I_i \left(1 + \frac{\Delta_{VTHyi}}{\eta V_T}\right) \tag{A.5}$$

Where the device is assumed to be in subthreshold operation, and ΔV_{THyi} is the threshold voltage mismatch between the reference and y^{th} output device. If we assume that the vector of currents from the input stack is

$$\vec{I} = \frac{g_m}{2^n} \sum_{x=1}^{n} \vec{k}_x V_x \tag{A.6}$$

where, for an n-input amplifier, $\vec{k}_x$ is the recombination vector, made of alternating strings of 1's and -1's with strings of length 2^{n-x}. This implies that the

x^{th} input will leak into the combined, y^{th} differential output current as

$$I_{outy} = \frac{g_m}{2^n} V_x \sum_{i=1}^{2^n} k_{yi} k_{xi} \left(1 + \frac{\Delta_{VTHyi}}{\eta V_T}\right) = \beta_{xy} g_m V_x \tag{A.7}$$

Once again, we can assume that the entries of $\vec{\Delta}_{VTHyi}$ are Gaussian random variables with zero mean and variance σ^2_{VTH}. For $x \neq y$, $\vec{k}_x$ and $\vec{k}_y$ are orthogonal, therefore the crosstalk term, β_{xy}, will be zero mean and have a variance of

$$\sigma^2_{\beta} = \left(\frac{1}{2^n \eta V_T}\right)^2 \sum_{i=1}^{2^n} k^2_{yi} k^2_{xi} \sigma^2_{VTHy} = \left(\frac{\sigma_{VTHy}}{\sqrt{2^n}\eta V_T}\right)^2 \tag{A.8}$$

BIBLIOGRAPHY

[1] E. Adrian. The electrical activity of the mammalian olfactory bulb. *Electroencephalogr Clin Neurophysiol*, 2:377–388, 1950.

[2] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walker. *Molecular Biology of The Cell*. Garland Science, New York, NY, 2002.

[3] C.A. Anastassiou, S. M. Montgomery, M. Barahona, G. Buzski, and C. Koch. The effect of spatially inhomogeneous extracellular electric fields on neurons. *J Neurosci*, 30:1925–1936, 2010.

[4] C. Andrews, L. Diamente, B. Johnson, and A. Molnar. A <12mW, 0.7-3.2ghz receiver with resonant multi-phase LO and current reuse harmonic rejection baseband. In *Radio Frequency Integrated Circuits Symposium (RFIC), 2012 IEEE*, pages 43–46, June 2012.

[5] C. Andrews, L. Diamente, D. Yang, B. Johnson, and A. Molnar. A wideband receiver with resonant multi-phase lo and current reuse harmonic rejection baseband. *IEEE J. Solid-State Circuits*, pages 1188–1198, May 2013.

[6] M.P. Brandon ans A.R. Bogaard, C.P. Libby, M.A. Connerney, K. Gupta, and M.E. Hasselmo. Reduction of theta rhythm dissociates grid cell spatial periodicity from directional tuning. *Science*, 332:595–599, 2011.

[7] J. Aziz, K. Abdelhalim, R. Shulyzki, R. Genov, B. Bardakjian, M. Derchansky, D. Serletis, and P. Carlen. 256-channel neural recording and delta compression microsystem with 3D electrodes. *IEEE J. Solid State Circuits*, pages 995–1005, March 2009.

[8] L. Berdondini, K. Imfeld, A. Maccione, M. Tedesco, S. Neukom, M. Koudelka-Hep, and S. Martinoia. Active pixel sensor array for high spatio-temporal resolution electrophysiological recordings from single cell to large scale neuronal networks. *Lab Chip*, pages 1–11, 2011.

[9] J. Beshel, N. Kopell, and L.M. Kay. Olfactory bulb gamma oscillations are enhanced with task demands. *J Neurosci*, 27:8358–8365, 2007.

[10] E. Boyden, F. Zhang, E. Bamberg, G. Nagel, and K. Deisseroth. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neuroscience*, pages 1263–1268, August 2005.

[11] J. Brunsilius, E. Siragusa, S. Kosic, F. Murden, E. Yetis, B. Luu, J. Bray, P. Brown, and A. Barlow. A 16b 80ms/s 100mw 77.6db snr cmos pipeline adc. In *Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2011 IEEE International*, pages 186–188, 2011.

[12] L. Buck and R. Axel. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell*, 65:175–187, 1991.

[13] S. Chang, K. AlAshmouny, M. McCormick, Y. Chen, and E. Yoon. BioBolt: a minimally-invasive neural interface for wireless epidural recording by intra-skin communication. In *VLSI Circuits (VLSIC), 2011 Symposium on*, pages 146–147, June 2011.

[14] T. A. Cleland. Early transformations in odor representation. *Trends Neurosci*, 33:130–139, 2010.

[15] T. A. Cleland and P. Sethupathy. Non-topographical contrast enhancement in the olfactory bulb. *BMC Neuroscience*, 7(7):1–18, Jan. 2006.

[16] M. Dandin, I. D. Jung, M. Piyasena, J. Gallagher, N. Nelson, M. Urdaneta, C. Artis, P. Abshire, and E. Smela. Post-CMOS packaging methods for integrated biosensors. In *Proc. of IEEE Sensors 2009*, pages 795–798, November 2009.

[17] S. V. David, N. Malaval, and S. A. Shamma. Decoupling action potential bias from cortical local field potentials. *Comput Intell Neurosci*, pages 1–12, November 2010.

[18] T. Delbruck and C.A. Mead. Adaptive photoreceptor with wide dynamic range. In *IEEE Int. Sym. on Circuits and Systems (ISCAS)*, volume 4, pages 339–342, may 1994.

[19] T. Denison, K. Consoer, W. Santa, A.-T. Avestruz, J. Cooley, and A. Kelly. A 2μW 100 nv/rthz chopper-stabilized instrumentation amplifier for chronic measurement of neural field potentials. *IEEE J. Solid-State Circuits*, 42(12):2934–2945, Dec. 2007.

[20] S. S. A. El-Rahman. Neuropathology of aluminum toxicity in rats (glutamate and GABA impairment). *Pharmacological Research*, pages 189–194, 2003.

[21] M. Elzeftawi, S. Beach, L. Wang, and L. Theogarajan. A 1.3μw 0.0075mm^2

neural amplifier and capacitor-integrated electrodes for high density neural implant recording. In *Biomedical Circuits and Systems Conference (BioCAS), 2013 IEEE*, pages 236–239, 2012.

[22] C. C. Enz, F. Krummenacher, and E.A. Vittoz. An analytical mos transistor model valid in all regions of operation and dedicated to low-voltage and low-current applications. *Analog Integrated Circuits and Signal Processing Journal on Low-Voltage and Low-Power Design*, 8(1):83114, July 1995.

[23] Q. Fan, J.H. Huijsin, and K.A.A. Makinwa. A 21nV/rthz chopper-stabilized multi-path current-feedback instrumentation amplifier with 2μV offset. *IEEE J. Solid-State Circuits*, pages 464–475, Feb. 2012.

[24] W. Franks, I. Schenker, P. Schmutz, and A. Hierlemann. Impedance characterization and modeling of electrodes for biomedical applications. *IEEE Trans. Biomed. Eng.*, pages 1295–1302, 2005.

[25] W.J. Freeman. Spatial properties of an eeg event in the olfactory bulb and cortex. *Electroencephalogr Clin Neurophysiol*, 44:586–605, 1978.

[26] U. Frey, U. Egert, D. Jackel, J. Sedivy, M. Ballini, P. Livi, F. Faraci, F. Heer, S. Hafizovic, B. Roscic, and A. Hierlemann. Depth recording capabilities of planar high-density microelectrode arrays. In *Neural Engineering, 2009. NER '09. 4th International IEEE/EMBS Conference on*, pages 207–210, May 2009.

[27] U. Frey, F. Heer, R. Pedron, M. Ballini, J. Mueller, D. Bakkum, S. Hafizovic, F. Faraci, F. Greve, K. Kirstein, and A. Hierlemann. Switch-matrix-based high density microelectrode array in CMOS technology. *IEEE J. Solid-State Circuits*, 45(2):467–482, Feb. 2010.

[28] D. Friedman and B.W. Strowbridge. Both electrical and chemical synapses mediate fast network oscillations in the olfactory bulb. *J Neurophysiol*, 89:2601–2610, 2003.

[29] P. Fries. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn Sci*, 9:474–480, 2005.

[30] R. F. Galn, N. Fourcaud-Trocm, G.B. Ermentrout, and N.N. Urban. Correlation-induced synchronization of oscillations in olfactory bulb neurons. *J Neurosci*, 26:3646–3655, 2006.

[31] H. Gao, R.M. Walker, P. Nuyujukian, K.A.A. Makinwa, K.V. Shenoy, B. Murmann, and T.H. Meng. HermesE: A 96-channel full data rate direct neural interface in 0.13 μm cmos. *IEEE J. Solid-State Circuits*, 47(4):1043–1055, Apr. 2012.

[32] Y. Gilad, O. Man, S. Paabo, and D. Lancet. Human specific loss of olfactory receptor genes. *Proc Natl Acad Sci*, 5(108):3324–3327, Feb. 2003.

[33] G. Girardeau, K. Benchenane, S.I. Wiener, G. Buzski, and M.B. Zugaro. Selective suppression of hippocampal ripples impairs spatial memory. *Nat Neurosci*, 12:1222–1223, 2009.

[34] D. H. Gire and N. E. Schoppa. Long-term enhancement of synchronized oscillations by adrenergic receptor activation in the olfactory bulb. *J Neurophysiol*, pages 2021–2025, February 2008.

[35] C.M. Gray, P. Knig, A.K. Engel, and W. Singer. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338:334–337, 1989.

[36] O. Gschwend, J. Beroud, and A. Carleton. Encoding odorant identity by spiking packets of rate-invariant neurons in awake mice. *PLoS ONE*, 7:e30155, 2012.

[37] Kwangseok Han, J. Gil, Seong-Sik Song, Jeonghu Han, Hyungcheol Shin, ChoongKi Kim, and Kwyro Lee. Complete high-frequency thermal noise modeling of short-channel mosfets and design of 5.2-ghz low noise amplifier. *IEEE J. Solid-State Circuits*, 40(3):726–735, March 2005.

[38] R. Harrison and C. Charles. A low-power low-noise CMOS amplifier for neural recording applications. *IEEE J. Solid-State Circuits*, 38(6):958–965, Jun. 2003.

[39] R. Harrison, P. Watkins, R. Kier, R. Lovejoy, D. Black, B. Greger, and F. Solzbacher. A low-power integrated circuit for a wireless 100-electrode neural recording system. *IEEE J. Solid-State Circuits*, 42(1):123–133, Jan. 2007.

[40] F. Heer, S. Hafizovic, T. Ugniwenko, U. Frey, W. Franks, E. Perriard, J.-C. Perriard, A. Blau, C. Ziegler, and A. Hierlemann. Single-chip microelectronic system to interface with living cells. *Biosens. Bioelectron.*, pages 2546–2553, May 2007.

[41] J. Holleman and B. Otis. A sub-microwatt low-noise amplifier for neural recording. In *IEEE 29th Annu. Int. Conf. Engineering Medicine and Biology Soc.*, pages 3930–3933, 2007.

[42] B. Johnson, D. DeTomaso, and A. Molnar. A low-power orthogonal current-reuse amplifier for parallel sensing applications. In *Proc. 2010 Eur. Solid-State Circuits Conf. (ESSCIRC)*, pages 318–321, Sept. 2010.

[43] B. Johnson, C. Lee, S. Sivaramakrishnan, and A. Molnar. A high-speed polar-symmetric imager for real-time calibration of rotational inertial sensors. In *Sensors, 2013 IEEE*, pages 1–4, 2013.

[44] B. Johnson and A. Molnar. A low-power orthogonal current-reuse amplifier for parallel sensing applications. *IEEE J. Solid-State Circuits*, 48(6):1487–1496, June 2013.

[45] B. Johnson, S. T. Peace, T. A. Cleland, and A. Molnar. A 50μm pitch, 1120-channel, 20khz frame rate microelectrode array for slice recording. In *Biomedical Circuits and Systems Conference (BioCAS), 2013 IEEE*, pages 109–112, 2013.

[46] B. Johnson, S. T. Peace, A. Wang, T. A. Cleland, and A. Molnar. Channel cmos microelectrode array with angle sensitive pixels for neuronal recording. *Sensors Journal*, 13(9):3211–3218, Sept. 2013.

[47] B. Johnson, S.T. Peace, T.A. Cleland, and A. Molnar. A scalable CMOS sensor array for neuronal recording and imaging. In *Proc. of IEEE Sensors 2011*, pages 924–927, Nov. 2011.

[48] J. B. Johnson. Thermal agitation of electricity in conductors. *Phys. Rev.*, 32(1):97–109, Dec. 1928.

[49] N. Joye, A. Schmid, and Y. Leblebici. A cell-electrode interface noise model for high-density microelectrode arrays. In *IEEE 31st Annu. Int. Conf. EMBS*, pages 3247–3250, September 2009.

[50] H. Kashiwadani, Y. F. Sasaki, N. Uchida, and K. Mori. Synchronized oscillatory discharges of mitral/tufted cells with different molecular receptive ranges in the rabbit olfactory bulb. *J Neurophysiol*, 82:1786–1792, 1999.

[51] L.M. Kay and W.J. Freeman. Bidirectional processing in the olfactory-limbic axis during olfactory behavior. *Behav Neurosci*, 112:541–553, 1998.

[52] K. Kotani and T. Ito. High efciency cmos rectier circuit with self vth-cancellation and power regulation functions for uhf rfids. In *Proc. IEEE ASSCC, Nov. 2007*, pages 119–122, 2007.

[53] A. Lambacher, V. Vitzthum, R. Zeitler, M. Eickenscheidt, B. Eversmann, R. Thewes, and P. Fromherz. Identifying firing mammalian neurons in networks with high-resolution multi-transistor array (MTA). *Appl Phys A*, pages 2613–2744, 2009.

[54] C. Lee, B. Johnson, and A. Molnar. A sub-threshold voltage ladder rectifier for orthogonal current-reuse neural amplifier. In *Biomedical Circuits and Systems Conference (BioCAS), 2013 IEEE*, pages 358–361, 2013.

[55] C. Lee and A. Molnar. Self-quenching, forward-bias-reset for single photon avalanche detectors in 1.8v, 0.18m process. In *Circuits and Systems (ISCAS), 2011 IEEE International Symposium on*, pages 2217–2220, 2011.

[56] T. Lee. *The Design of CMOS Radio-Frequency Integrated Circuits, Second Edition*. Cambridge University Press, Cambridge, UK, 2003.

[57] H. Linden, T. Tetzlaff, T. Potjans, K. Pettersen, S. Grun, M. Diesmann, and G. T. Einevoll. Modeling the spatial reach of the LFP. *Neuron*, 72(5):859–872, Dec. 2011.

[58] C. Linster and T.A. Cleland. Decorrelation of odor representations via spike timing-dependent plasticity. *Front Comput Neurosci*, 4:157, 2010.

[59] W. Liu and Z. Yang. Engineering hope with biomimetic microelectronic systems. In *Proc. 2010 Eur. Solid-State Circuits Conf. (ESSCIRC)*, pages 17–26, Sept. 2010.

[60] Y. Liu, E. Smela, N. Nelson, and P. Abshire. Cell-lab on a chip: a CMOS-based microsystem for culturing and monitoring cells. In *IEEE 26th Annu. Int. Conf. EMBS*, pages 2534–2537, September 2004.

[61] C. M. Lopez, A. Andrei, S. Mitra, M. Welkenhuysen, W. Eberle, C. Bartic, R. Puers, R. F. Yazicioglu, and G. Gielen. An implantable 455-active-electrode 52-channel cmos neural probe. *IEEE J. Solid-State Circuits*, 49(1):1–14, Jan 2014.

[62] Bazhenov M, Rulkov N.F., and Timofeev I. Effect of synaptic connectivity

on long-range synchronization of fast cortical oscillations. *J Neurophysiol*, 100:1562–1575, 2008.

[63] K. MacLeod, A. Bcker, and G. Laurent. Who reads temporal information contained across synchronized and oscillatory spike trains? *Nature*, 395:693–698, 1998.

[64] V. Majidzadeh, A. Schmid, and Y. Leblebici. Energy efficient low-noise neural recording amplifier with enhanced noise efficiency factor. *IEEE Trans. Biomed. Circuits Syst.*, 5(3):262–271, Jun. 2011.

[65] C. Mead. *Analog VLSI and Neural Systems*. Addison-Wesley, Reading, MA, 1989.

[66] A. Molnar, B. Lu, S. Lanzisera, B.W. Cook, and K.S.J. Pister. An ultra-low power 900MHz RF transceiver for wireless sensor networks. In *IEEE Custom Integrated Circuits Conf.*, pages 401–404, Oct. 2004.

[67] R. Muller, S. Gambini, and J. Rabaey. A 0.013mm^2 5 μW DC-coupled neural signal acquisition IC with 0.5V supply. *IEEE J. Solid-State Circuits*, pages 232–243, Jan. 2012.

[68] W. Muller and J. Connor. Dendritic spines as individual neuronal compartments for synaptic ca^{2+} responses. *Nature*, (354):73–76, Nov. 1991.

[69] M. Pelgrom, H. Tuinhout, and M. Vertregt. Transistor matching in analog cmos applications. In *IEDM Dig. of Tech. Papers*, pages 915–918, Dec. 1998.

[70] C. Peters, J. Handwerker, D. Maurath, and Y. Manoli. A sub-500 mv highly efficient active rectifier for energy harvesting applications. *IEEE Trans. Circuits Syst.*, pages 1542–1550, July 2011.

[71] C. Ponn and K. Zhou. Neuromorphic silicon neurons and large-scale neural networks: challenges and opportunities. *Front Neurosci*, 5(108):1–3, Sept. 2011.

[72] C. Qian, J. Parramon, and E. Sanchez-Sinencio. A micropower low-noise neural recording front-end circuit for epileptic seizure detection. *IEEE J. Solid-State Circuits*, 46(6):1392–1405, Jun. 2011.

[73] R. Sarpeshkar. *Ultra Low Power Bioelectronics*. Cambridge University Press, Cambridge, UK, 2010.

[74] N. E. Schoppa. Synchronization of olfactory bulb mitral cells by precisely timed inhibitory inputs. *Neuron*, 49:271–283, 2006.

[75] I. Shcherback, A. Belenky, and O. Yadid-Pecht. Empirical dark current modeling for complementary metal oxide semiconductor active pixel sensor. *Opt. Eng.*, 41:1216–1219, Jan 2002.

[76] J. Shi, J. Pandey, and B. Otis. A 1.5GHz high-Q receiver based on current reuse. In *IEEE Int. Sym. on Circuits and Systems (ISCAS)*, pages 1800–1803, May 2011.

[77] D. B. Shire, S. K. Kelly, J. Chen, P. Doyle, M. D. Gingerich, S. F. Cogan, W. A. Drohan, O. Mendoza, L. Theogarajan, J. L. Wyatt, and J. F. Rizzo. Development and implantation of a minimally invasive wireless subretinal neurostimulator. *IEEE Trans. Biomed. Eng.*, 56(10):2502–2511, Oct. 2009.

[78] I. SM. Steyaert, W. Sansen, and C. Zhongyuan. A micropower low-noise monolithic instrumentation amplifier for medical purposes. *IEEE J. Solid-State Circuits*, 22(6):1163–1168, Dec. 1987.

[79] M. Stopfer, S. Bhagavan, B.H. Smith, and G. Laurent. Impaired odour discrimination on desynchronization of odour-encoding neural assemblies. *Nature*, 390:70–74, 1997.

[80] P.T. Theilmann, C.D. Presti, D.J. Kelly, and P. Asbeck. A μw complementary bridge rectifier with near zero turn-on voltage in sos cmos for wireless power supplies. *IEEE Trans. Circuits Syst.*, pages 2111–2124, Sept. 2012.

[81] L. S. Theogarajan. A low-power fully implantable 15-channel retinal stimulator chip. *IEEE J. Solid-State Circuits*, 43(10):2322–2337, Oct. 2008.

[82] R.D. Traub, M.A. Whittington, S.B. Colling, G. Buzski, and J.G. Jefferys. Analysis of gamma rhythms in the rat hippocampus in vitro and in vivo. *J Physiol (Lond)*, 2:471–484, 1996.

[83] A. A. Trusov, S. A. Zotov, B. R. Simon, and A. M. Shkel. Silicon accelerometer with differential frequency modulation and continuous self-calibration. In *Micro Electro Mechanical Systems (MEMS), 2013 IEEE 26th International Conference on*, pages 29–32, 2013.

[84] A. Wang, P. R. Gill, and A. Molnar. Angle sensitive pixels in CMOS for lensless 3D imaging. In *Proc. of IEEE CICC*, pages 371–374, November 2009.

[85] A. Wang, P. R. Gill, and A. Molnar. Fluorescent imaging and localization with angle sensitive pixel arrays in standard CMOS. In *Proc. of IEEE Sensors 2010*, pages 1706–1709, November 2010.

[86] A. Wang and A. Molnar. A light field image sensor in 180 nm CMOS. *IEEE J. Solid-State Circuits*, 47(1):257–271, Jan 2012.

[87] W. Wattanapanitch, M. Fee, and R. Sarpeshkar. An energy-efficient micropower neural recording amplifier. *IEEE Trans. Biomed. Circuits Syst.*, 1(2):136–147, Jun. 2007.

[88] M.A. Whittington, R.D. Traub, N. Kopell, B. Ermentrout, and E.H. Buhl. Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int J Psychophysiol*, 38:315–336, 2000.

[89] R. Wodnicki, G.W. Roberts, and M. D. Levine. A foveated image sensor in standard cmos technology. In *Custom Integrated Circuits Conference, 1995*, pages 357–360, 1995.

[90] Z. Yang, Q. Zhao, E. Keefer, and W. Liu. Noise characterization, modeling, and reduction for in vivo neural recording. *Advances in Neural Information Processing Systems*, pages 2160–2168, 2010.

[91] M. Yin and M. Ghovanloo. A flexible clockless 32-ch simultaneous wireless neural recording system with adjustable resolution. In *ISSCC Dig. Tech. Papers*, pages 432–433, 2009.

[92] T. P. Zanos, P. J. Mineault, and C. C. Pack. Removal of spurious correlations between spikes and local field potentials. *J Neurophysiol*, pages 474–486, January 2011.

[93] F. Zhang, J. Holleman, and B. P. Otis. Design of ultra-low power biopotential amplifiers for biosignal acquisition applications. *IEEE Trans. Biomed. Circuits Syst.*, 6(4):344–355, Aug. 2012.

[94] A. Zolfaghari and B. Razavi. A low-power 2.4-GHz transmitter/receiver CMOS IC. *IEEE J. Solid-State Circuits*, 38(2):176–183, Feb. 2003.